Deciding to consult the general practitioner for joint pain: a choice-based conjoint analysis study

Domenica Coxon

A thesis submitted for the degree of Doctor of Philosophy

June 2013

Arthritis Research UK Primary Care Centre, Keele University

Declaration

This PhD involved the collection of primary data and this was funded by an Arthritis Research UK Programme grant (18174) obtained at the Arthritis Research UK Primary Care Centre by Professor George Peat and colleagues.

Throughout the PhD project, with guidance from my supervisors Professor George Peat, Dr Martin Frisher, Dr Clare Jinks and statistical advisor Dr Kelvin Jordan, I developed the questionnaire, gained ethical approval and coordinated and managed it alongside the thesis. In this thesis, I have generated primary research and had primary control over the content of the survey, with the guidance of supervisors. The sample frame for this PhD was taken from the North Staffordshire Osteoarthritis Project (NorStOP). I conducted the data cleaning of the primary data collected for the project. Professor George Peat and Dr Clare Jinks advised on the procedures used to collect data and measures taken to enhance response rates.

Professor George Peat, Dr Martin Frisher and Dr Clare Jinks advised me on the planning of all analyses and on the writing and presentation of chapters. I conducted all analyses and wrote the chapters myself. I received guidance on the search strategy for the systematic review from Professor George Peat, Dr Martin Frisher and Dr Clare Jinks. Dr Kelvin Jordan provided statistical advice on the conjoint design and statistical analysis.

Abstract

A substantial proportion of older adults with non-inflammatory joint pain do not consult their general practitioner (GP) despite apparent clinical need. This thesis describes the development, execution, and interpretation of an original study using conjoint analysis – a fairly novel approach with some advantages over conventional observational and qualitative studies - to understand the relative importance of need-related and service-related factors on the decision to consult the GP.

Background reading, a systematic review of previously published conjoint analysis studies, and a series of developmental studies involving patients and members of the public informed the design of the main study.

A partial-profile choice-based conjoint (PPCBC) questionnaire was chosen, comprising 10 choice tasks using a combination of selected attributes (pain characteristics, pain disruption to everyday life, comorbidity, assessment and investigations available, available treatment options, and perceived GP attitude). The PPCBC questionnaire was postally-administered to 1170 adults aged 50 years and over with hip, knee, or hand pain identified from an existing population cohort study in North Staffordshire.

863 questionnaires were returned (adjusted response rate 74%; mean age: 70 years; 55% female) and well-completed (<5% missing data). The extent to which pain disrupted everyday life (1.10 logits) and perceived GP attitude (0.86 logits) were the most important determinants of the decision to consult the GP. Service factors were highly influential with a 'negative' GP attitude potentially outweighing the perceived value of optimal assessment and management. Latent class analysis identified possible subgroups with differing strengths of preference.

Conjoint analysis is feasible and offers unique insights into the relative importance of actual and hypothetical services. While it presents many challenges - extensive developmental testing, complex design and analysis procedures, ability to integrate findings from a range of different methods – it can provide important information on patients' preferences for existing and emerging treatments and models of care.

Acknowledgements

Firstly, I would like to thank my supervisors Professor George Peat, Dr Martin Frisher and Dr Clare Jinks for their invaluable support and guidance on the drafts of this thesis. I would also like to thank my statistical advisor (Dr Kelvin Jordan) and the other members of the project team (including Ian Thomas, Charlotte Clements, Jo Bailey, Jacqueline Gray and Rhian Hughes). I would also like to thank Basem Al-Omari for acting as a second independent reviewer in the systematic review. A million thanks also to the members of the Research Users' Group at the Arthritis Research UK Primary Care Centre, and most importantly to those people who responded to the conjoint questionnaire. Many thanks also to Claire Ashmore for transcription of the cognitive interviews.

My PhD studentship and research project was funded jointly by the Arthritis Research UK Primary Care Research Centre and the Faculty of Health at Keele University. I would also like to acknowledge the Psychology Postgraduate Affairs Group for awarding a conference bursary for attendance at the Society for Academic Primary Care Northern conference and the British Society for Rheumatology national conference. Thanks to all the staff and students at Arthritis Research UK Primary Care Centre for their encouragement and support.

In addition, I would like to thank the advisors at SKIMgroup and Sawtooth Software (Inc. Orem, UT), in particular Stefan Ammerlaan, Brian McEwan and Jefferson Faulkner for their helpful clarifications and technical support. Finally, I would like to express my thanks to Raymond Sharpe, my friends, my cat Oscar James and my family; past, present and future for their enduring love and support.

Context of the thesis

I obtained a BA Psychology Major (with Sociology) from the University of Northampton in 2006. This inspired me to do a Postgraduate Diploma (Year 1 certificate) in psychotherapy and counselling psychology. However, before I enter into the clinical psychology field I wanted to conduct my own research (with a strong focus on psychology) and undertake a PhD using a novel methodology to analyse the internal processes of decision-making. I successfully applied for a three-year full-time PhD studentship advertised at the Arthritis Research UK Primary Care Centre. The challenge was to convert a rather broad and speculative remit into a focused and feasible PhD project.

Table of contents

List of tables

8 Chapter Eight: Results II: direct rating of selected attributes, conjoint utilities and relative importance of attributes

9 Chapter Nine: Results III: subgroup analysis

10 Chapter Ten: Discussion, conclusions and recommendations

List of figures

List of abbreviations

ACA	Adaptive Conjoint Analysis
ACBC	Adaptive Choice-Based Conjoint
ARUKPCC	Arthritis research UK Primary Care Centre
CA	Conjoint Analysis
CBC	Choice-Based Conjoint
CI	Cognitive Interviews
CI	Confidence Interval
CSQ	Coping Strategies Questionnaire
CVA	Conjoint value analysis
DCE	Discrete Choice Experiment
GP	General Practitioner
HADS	Hospital Anxiety and Depression Scale
IPQ-R	Illness Perceptions Questionnaire-Revised
IQR	Inter-quartile Range
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
MNL	Multinomial Logistic Regression
MSF	Musculoskeletal Services Framework
NHS	National Health Services
NICE	National Institute of Clinical Excellence
NorStOP	North Staffordshire Osteoarthritis Project
NRS	Numerical Rating Scale
OA	Osteoarthritis
OR	Odds Ratio

PPCBC	Partial-Profile Choice-Based Conjoint
REC	Research Ethics committee
RP	Revealed Preference methods
RUG	Research User Group
R&D	Research and Development
SD	Standard Deviation
SE	Standard Error
SP	Stated Preference methods
UK	United Kingdom
US(A)	United States (of America)
6YFU	Six year Follow-UP

Publications and presentations arising from this thesis

Peer reviewed publications:

Coxon D, Frisher M, Jinks C, Jordan K, Peat G. 2012. Deciding to consult the general practitioner for symptomatic osteoarthritis: a choice-based conjoint analysis study, Rheumatology (vol. 51. P. 37) **(Abstract)**

Oral presentations:

Coxon D, Frisher M, Jinks C, Jordan K, Peat G. Deciding to consult the general practitioner for joint pain: a choice-based conjoint analysis study.

British Society for Rheumatology. Glasgow, Scotland. 1– 3 May 2012

Coxon D, Frisher M, Jinks C, Jordan K, Peat G. The decision to consult the general practitioner for joint pain: a choice-based conjoint analysis study.

Society for Academic Primary Care. Kendal, England. 24– 25 November 2011

Poster presentations:

Coxon D, Frisher M, Jinks C, Jordan K, Peat G. Deciding to consult the general practitioner for joint pain: a choice-based conjoint analysis study.

North American Primary Care Research Group. New Orleans, United States of America. December 2012

See Appendix 1a for full abstracts of the above.

1 Chapter One: Introduction and background to the determinants of general practice consultation for joint pain in older adults

1.1 Introduction

The primary aim of this thesis is to quantify the relative importance of selected clinical need and general practice service factors ('attributes') in the decision to consult the general practitioner (GP) in older adults with joint pain/osteoarthritis. This involves the application of a novel method called conjoint analysis (CA). CA works by asking respondents to choose between alternative hypothetical scenarios to decide which would most likely lead them to consult a GP for joint pain. Respondents' hypothetical choices are then used to deduce their priorities for GP consultation and the trade-offs made between different attributes presented at different levels. This is preferable to traditional methods (which ask respondents to rate directly how important certain attributes are) because it forces respondents to make difficult choices and trade-offs. This enables researchers to reveal and quantify respondents underlying preferences more precisely.

This chapter begins with a brief overview of the occurrence of joint pain in the community (section 1.2) followed by the prevalence of joint pain in primary care (1.3). The next section outlines why it is important to investigate the determinants of consultation for joint pain (section 1.4), before considering in more detail what is already known about the factors associated with consulting general practice for joint pain (1.5). The final section considers the methodological limitations of previous research (section 1.6), before introducing conjoint analysis in the next chapter (chapter 2), as a potentially useful method for understanding the determinants of consultation.

1.2 The occurrence of joint pain/ symptomatic osteoarthritis in the community

The majority of peripheral joint pain in older adults is likely to be attributed to osteoarthritis (OA) (Zhang et al. 2010a). OA refers to *'a clinical syndrome of joint pain accompanied by varying degrees of functional limitation and reduced quality of life'* (National collaborating centre for chronic conditions 2008, p.3). The knees, hips and small joints of the hands are most frequently affected. It has been estimated that 10% of the global population aged 60 years and over have significant clinical problems that can be attributed to OA (World Health Organization 2003). OA has thus been identified as a major condition dominating the overall burden of non-fatal, disabling conditions and a major contributor to years lived with disability (Mathers et al. 2002).

OA is one of the most common causes of disability in people aged 50 and above, especially in women (National collaborating centre for chronic conditions 2008). Official government statistics indicate that by 2026 the number of persons aged 60 and above in Britain will reach 17.1 million (Office of National Statistics 1996), an increase of almost fifty percent on the same age category in 1996. As a consequence of the growing proportion of older adults in the population, the numbers of people at risk for disorders associated with ageing, (e.g. hip OA), are projected to increase (Birrell et al. 1999, Perruccio et al. 2006). One study suggests that one in four people may develop symptomatic hip OA in their lifetime (Murphy et al. 2010). The number of hip replacements in the UK is subsequently set to increase by 40% over the next 30 year period (Birrell et al. 1999).

1.3 Consultation prevalence of joint pain/OA in primary care and clinical guidelines for OA treatments

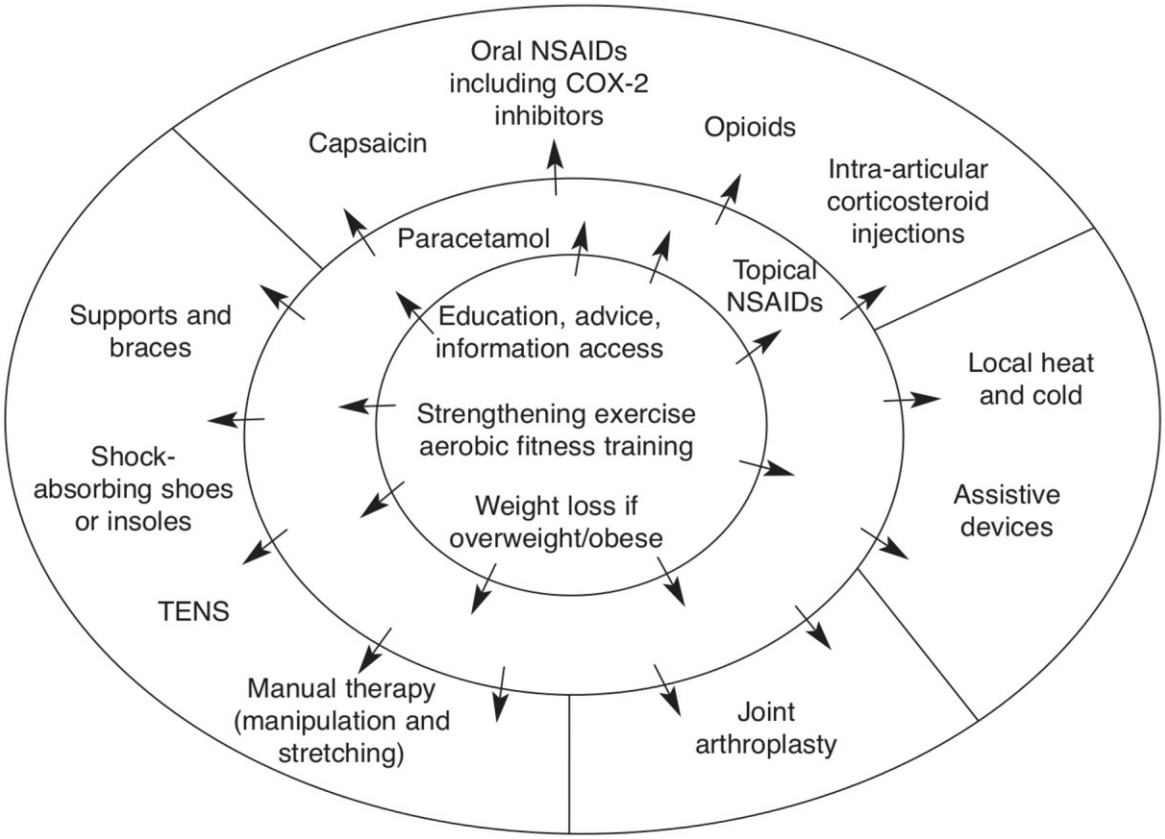
In the UK, general practice provides the main first point of contact with formal healthcare services and accounts for an estimated 90% of patient contacts in the National Health Service (NHS) (Mant 1998). OA is one of the most common diagnoses made in general practice in older adults (McCormick et al. 1995) and primary care represents the setting for ongoing care for older adults with symptomatic OA (National collaborating centre for chronic conditions 2008). The overall persons consulting prevalence rates for OA derived from two UK general practice databases in one year is 230–280 per 10,000 people aged 15 years and over (Jordan et al. 2007). In one database for example, the consulting prevalence rate for OA increased markedly from a rate of 320 per 10,000 people aged 45-64 years up to a rate of 1050 per 10,000 people aged 75 years and over. Females have a higher consulting prevalence than men for OA (Jordan et al. 2007).

However, recent studies suggest that, in common with many other symptoms in the general population (Green et al. 2001) only a minority of older people with painful joints will consult their general practitioner about it in the course of any one year (Jinks et al. 2004, Bedson et al. 2007). For example, while the population prevalence of knee pain lasting at least 4 weeks in the past year in adults aged 55 years and over is estimated to be 25% (Peat et al. 2001), the annual person-consulting consultation prevalence for any knee disorders in patients aged 45 years and over is estimated to be between 4 to 7% (Jordan et al. 2010). This may underestimate the actual consultation rate due to non-coding or use of generalised codes (e.g. generalised osteoarthritis). In studies that have specifically investigated consultation (reported or recorded) within the same population sample, it appears that roughly one third of older adults with knee pain have

consulted their GP in the past year (Jinks et al. 2004, Jordan et al. 2006). While mild or transient symptoms account for many non-presented cases of joint pain, Bedson et al. (2007) found that only 50% of older adults with severe knee pain and disability had consulted their GP about it in the previous 18 months.

Contrary to the belief that 'nothing can be done' for osteoarthritis, recent national (National collaborating centre for chronic conditions 2008), European (Zhang et al. 2010a) and international (Zhang et al. 2010b) guidelines recommend a wide range of effective and cost-effective pharmacological, nonpharmacological and surgical interventions for OA. The interventions recommended by NICE are shown in Figure 1.1 (see page 5). In the UK, these treatments are available over the counter, from the multidisciplinary primary care team or the GP acts as gatekeeper for access to specialist providers. Clearly, core interventions – weight loss if overweight and exercise – may confer health benefits beyond just the control of OA pain and maintenance of function (e.g. reduction of cardiovascular risk, management of depression, falls prevention). Arguably, the successful adoption of weight loss and exercise in a large section of the population for the management of OA would serve a public health function (Hartvigsen & Christensen 2007). However, there has been repeated evidence from surveys in the UK and elsewhere that many of these effective interventions are under-used in the management of OA (Porcheret et al. 2007, Steel et al. 2008, Peat & Thomas 2009, Li et al. 2011).

Figure 1.1: NICE clinical guidelines for osteoarthritis treatments (National collaborating centre for chronic conditions 2008)



1.4 Why it is important to investigate determinants of consultation in people with joint pain/OA

The importance of investigating the determinants of nonconsultation in people with joint pain/OA depends to some extent on one's interpretation of the meaning and significance of nonconsultation. It can be considered as a positive phenomenon, i.e. an indication of successful adaptability and self-management in the face of the physical and emotional challenges of joint pain (Huber et al. 2011) or as a negative phenomenon, i.e. an indication that there are people not accessing healthcare that can prevent acute problems becoming chronic (Hartvigsen & Christensen 2007).

There is evidence to suggest that repeated consultation is deemed to be of limited value to joint pain patients because no effective treatment can be offered

(Jinks et al. 2007). For example, older adults with knee pain continue to have persistent problems irrespective of whether they have consulted their GP or not (Blagojevic et al. 2008). Moreover, nonconsulters who report high physical disability frequently give positive reasons for not consulting, such as perceived emotional and practical benefits like independence from the health system and personal autonomy (Ong et al. 1999).

In other words in the context of a definition of health that is based on an individuals' ability to *'adapt and self-manage in the face of social, physical, and emotional challenges'* (Huber et al. 2011, p.1) nonconsultation might be considered a positive phenomenon. Nonconsultation might indicate a successful development of Huber's 'social' domain of health, which is the ability for patients to feel healthy and self-manage their condition by pursuing independence in their social environment (e.g. continuing to work) (Huber et al. 2011). The work by Lorig et al (1999) suggests that enabling patients to self-manage their chronic health problem better can lead to a reduction in distress, social limitations and healthcare costs in addition to improving self-rated health (Lorig et al. 1999). From the perspective of the National Health Service (NHS) it is important to acknowledge that nonconsultation may be viewed as a positive phenomenon, indicative of better self-management and adaptability and reducing the danger of swamping an already costly and busy health service.

There are also concerns that there is an over emphasis on the use of prescription drugs in modifying OA and that this leads to over "medicalisation" of people with OA (Dieppe 2005). Nonconsultation thus may be considered a positive antidote to such medicalisation.

However, balanced against these considerations for nonconsultation as a positive phenomenon, the case for considering nonconsultation as a negative phenomenon can also be made. For example, because up to half of those with severely disabling joint pain may not visit their GP about it over a prolonged period (Bedson et al. 2007) it might be assumed that there is a level of unmet need. A level of unmet need might be assumed because these patients are not accessing a wide range of support and treatment potentially available in primary care. There is evidence of unmet need by patients with knee OA for information and support with strategies for coping with OA and maintaining independence (Victor et al. 2004). It can be argued that consultation behaviour driven by episodes of severe disruption is a fairly ineffective approach to establishing the longer-term changes in lifestyle (e.g. weight reduction, exercise/physical activity, joint protection) that are needed to reduce the risk of progression over time (Hartvigsen & Christensen 2007). Nonconsultation is a significant issue for general practice because early intervention (in knee pain for example) such as advice about weight loss, exercises, physiotherapy and basic analgesia, may improve the long-term prognosis for these patients (Deyle et al. 2000). Thus it is the delayed support and help with self-management for nonconsulters with severe problems that may have more serious consequences (Bedson et al. 2007). According to McDonald et al. (2008), OA patients that under-report pain to a health professional are more likely to receive inadequate assessment and management advice (McDonald et al. 2008) and thus suffer more pain.

In general the under-treatment of pain in the elderly (Berry et al. 2001) suggests that a more proactive approach is needed in general practice to manage (knee) pain for nonconsulters (Bedson et al. 2007). In the main, more research

into the consequences of nonconsultation for joint pain/OA is needed (Bedson et al. 2007).

Overall, the emphasis in this thesis is not on the objective of simply increasing the frequency of GP consultation for joint pain. If older adults with severe pain decide not to consult a GP for it, it does not necessarily indicate unmet need for healthcare. Nonconsultation may indicate successful adaptation and self-management (Huber et al. 2011) and the exercise of autonomy. However, when considering the impact of severity of pain upon nonconsulters the emphasis is upon missed opportunities for lifestyle, self-management advice and secondary prevention (as well as factors like information-delivery skills) that can be offered in GP consultations.

1.4.1 Defining healthcare 'need' in the joint pain/symptomatic OA population

The importance of healthcare needs for musculoskeletal conditions was highlighted in UK healthcare policy by the introduction of the Musculoskeletal Services Framework (MSF) (Department of Health 2006). The MSF framework sought to provide advice, assessment and treatment to enable people with musculoskeletal conditions to maintain independence and increase health potentials (Department of Health 2006). Its central goal is to address unmet needs by focusing on the broader context of needs assessment for musculoskeletal conditions (Department of Health 2006).

Currently, the dominant decision-making framework for health needs assessment defines 'need' as *'the ability to benefit from healthcare, which depends both on morbidity and on the effectiveness of care'* (Stevens & Gabbay 1991). Needs assessment frameworks involve identifying 'felt need' (an

individuals' personal assessment that they have need for healthcare) and an investigation of whether 'felt needs' are turned into an 'expressed' need (demand for healthcare – i.e. when patients consult) (Bradshaw 1994).

Jinks et al. (2007) highlight that the belief that knee pain is part of normal ageing and that few effective treatments are available may prevent 'felt' needs becoming 'expressed' needs in adults with knee pain. In adults with knee pain there was evidence of missed opportunities for effective interventions (e.g. lifestyle advice) (Jinks et al. 2007). This thesis places its focus on the people with unexpressed and unmet needs who may benefit from consultation for their joint pain. Another explanation for unexpressed need is the influence of competing comorbidities. Patients who present to general practice with musculoskeletal symptoms and conditions often have other health problems (Jinks et al. 2007). One qualitative study found that many older people who lived with more than one condition ranked them in terms of severity and perceived urgency (Jinks et al. 2007). Thus 'felt' need for healthcare for joint pain may not be 'expressed' because the long list of other illnesses may push it off the patients' agenda for consultation. It may be that there is a need to engage with the older population to shift the perception that 'nothing can be done' for their joint pain, something that may be getting in the way of 'felt' need being translated into 'expressed' need for care.

Overall, an understanding of how 'felt' need might become translated into 'expressed' need (i.e. in the consultation room) can be used by healthcare policy-makers and clinicians in order to improve primary care needs assessment and prevent missed opportunities for healthcare for people with musculoskeletal conditions. Understanding the reasons for not expressing need may encourage

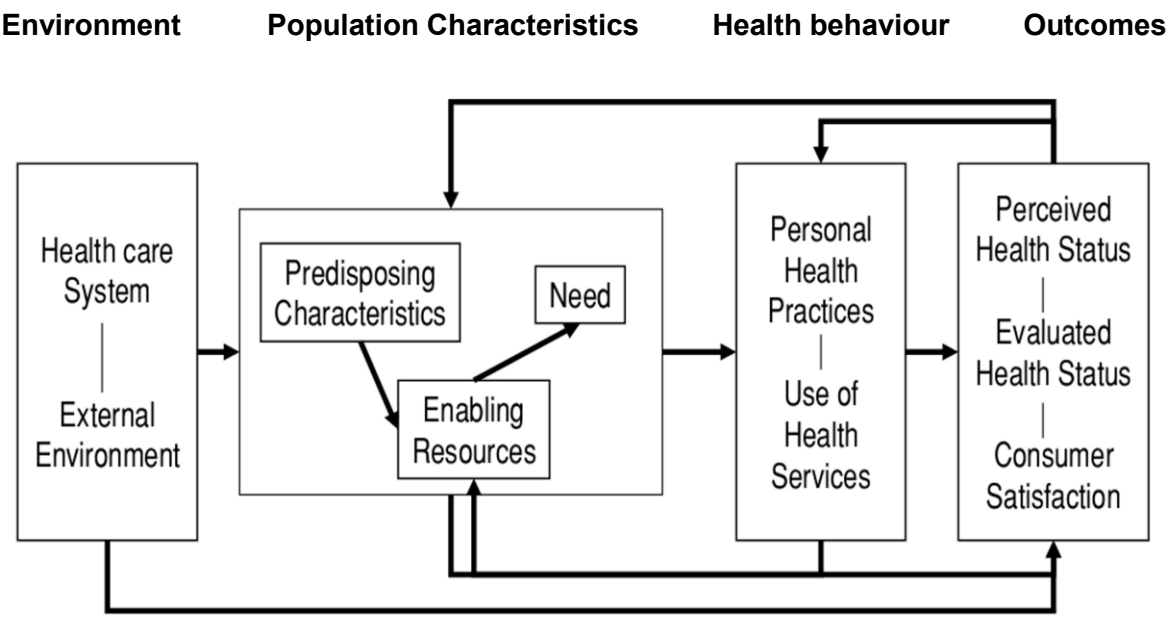
clinicians to inform patients about the newly emerging range of evidence about effective interventions that are currently available (National collaborating centre for chronic conditions 2008, Roddy et al. 2005) but are being under-used in the management of OA (Porcheret et al. 2007, Steel et al. 2008, Peat & Thomas 2009, Li et al. 2011). This might lead to patients starting to believe that something (rather than nothing) can be done for their joint pain.

1.5 Understanding the determinants of consultation

1.5.1 The Andersen-Newman model of healthcare utilisation

The Andersen-Newman model of healthcare utilisation (HCU) (Andersen 1995) developed by Andersen and colleagues in 1973 and updated in 1995, is a biopsychosocial framework. It proposes that HCU is determined by individual, societal and health service factors. Individual factors are organised into three categories: need, enabling and predisposing (see Figure 1.2).

Figure 1.2: The Andersen-Newman model of healthcare utilisation revisited, developed by Andersen et al. (1995)



Need factors include an individuals evaluated and perceived functional ability, symptoms, and general health status (Andersen 1995). These are based on ‘expressed’ need and may not capture aspects of ‘felt’ need. Enabling factors include accessibility of healthcare services. Predisposing factors include age, gender, education and occupation, in addition to beliefs about disease and attitudes toward health services (Andersen 1995).

The Andersen–Newman model (Andersen 1995) is one of the most widely used frameworks used to analyse the factors that are associated with healthcare utilisation. A systematic review by Phillips et al. (1998) identified 139 published studies on formal HCU that had specifically stated their use of the Andersen–Newman model as a conceptual framework. An expert panel researching factors associated with help-seeking from a medical gatekeeper for knee pain (Dieppe et al. 1999) proposed that healthcare seeking could be understood better when incorporating all categories of determinants used in the Andersen-Newman model. Several primary studies of the determinants of primary care consultation for osteoarthritis in the United Kingdom (Dieppe et al. 1999, Jordan et al. 2006, Bedson et al. 2007) and other countries (Cronan et al. 1995, Kim et al. 2010) have either explicitly used this model or can be evaluated using it.

Despite its popularity, the Andersen-Newman model (Andersen 1995) has been criticised in two ways. Firstly, although it identifies a range of factors that may influence HCU it does not explore how these factors interrelate (Bradley et al. 2002). For example, knowledge of services may interrelate with attitude towards care providers (Bradley et al. 2002) and failing to recognise these interrelationships may oversimplify the role of some factors in HCU (Phillips et al. 1998).

Secondly, despite being specified in the predisposing factors in the original Andersen-Newman model outline (1995) the model has been criticised for over-emphasising individual factors and ignoring the role of system delivery factors (De Boer et al. 1997). Attitudes toward health services are commonly absent from research into elderly populations (Strain 1991). The relative importance of expectations and attitudes towards health services for the elderly population is not clear (Strain 1991). The role of characteristics of the physician upon HCU requires deeper investigation (de Boer et al. 1997). Bradley et al. (2002) expanded the Andersen-Newman framework to include specific psychosocial factors (including attitudes towards the care providers; their technical expertise and interpersonal skills). Bradley et al. (2002) recognised them as key determinants of service that are absent and argued that an identification of them might enhance our understanding of HCU (Bradley et al. 2002).

1.5.2 Previous studies of determinants of general practice consultation for joint pain/symptomatic OA

De Boer et al. (1997) conducted a systematic review of published studies on determinants of HCU in adults with chronic illness. Based on 53 studies into hospitalizations and physician visits they concluded that psychological factors (psychological distress and depression) and 'need factors' (e.g. disease severity) were the most important factors. The relationship between physician visits and disease severity was less evident. Predisposing factors (e.g. age and sex) and enabling factors (e.g. income and social support) appeared not to influence HCU. Disease duration and comorbidity appeared to have no effect but quality of life, perceived health and aspects of the hospitals may influence HCU.

Factors associated with healthcare use (non-UK) for joint pain

The following section reviews six studies which include factors associated with healthcare use (non-UK) for joint pain and six studies which include factors associated with GP consultation (UK) for joint pain specifically, many of which post-date the De Boer et al. (1997) review. The question of whether the non-UK studies (see Table 1.1) are representative of determinants of consultation in the UK must be taken into account.

Table 1.1: Factors associated with healthcare use (non-UK) for joint pain

Andersen-Newman (1995) framework factors	Healthcare use (non-UK)
Predisposing	Age (Cronan et al. 1995, Rao et al. 1997) Male (Rao et al. 1997) Single (Rosemann et al. 2007) Health insurance (Rao et al. 1997)
Need-related	Impairment (Cronan et al. 1995) Well-being (Cronan et al. 1995) Pain chronicity (Hopman-Rock et al. 1997) HRQOL (Ethgen et al. 2002, Dominick et al. 2004) Mood (Rosemann et al. 2007) Pain severity (Rosemann et al. 2007) Pain frequency (Dominick et al. 2004) Drug prescriptions (Rosemann et al. 2007) Health status (Rao et al. 1997) Activity/work limitations due to arthritis (Rao et al. 1997) Overweight (Rao et al. 1997)
Enabling	Prior use of the health system (Cronan et al. 1995) Attendance of a physiotherapist (Hopman-Rock et al. 1997)

Cronan et al. (1995) and Rao et al. (1997) found that older age was associated with increased healthcare use. Although, Cronan et al. (1995) found that older age, greater impairment, and a lower quality of well-being were predictors of HCU, the strongest predictor was prior usage of the healthcare system in HCU (in the past 12 months) in Health Maintenance Organisations (HMOs) members aged 60 and above with OA in the USA. A model including prior

HCU, age, quality of well-being, physical impairment and pain accounted for 29.5% of the variance in HCU rates. This study finding contradicts the finding in De boer et al. (1997) review that age was not a consistent predictor of HCU. Generalisability to the population at large was limited since only 25% of the population agreed to participate – raising issues of non-response bias. Furthermore, respondents were predominantly caucasian, married and with some education and as members of HMOs they were charged only a small fee for healthcare compared to those with other healthcare access (raising issues of selective non-response).

Rao et al. (1997) found that patients who were male, younger, had better self-perceived health, were not overweight, had fewer activity/work limitations and no health insurance were less likely to consult a doctor for arthritis. Rao et al. (1997) compared consulters and nonconsulters for self-reported arthritis within a stratified random sample in the U.S. Identified from cross-sectional data from the 1989 National Health Interview Survey (NHIS) of those who reported arthritis (n = 2944), 16.4% reported not consulting a doctor for arthritis. Of those who had never consulted a doctor for arthritis, 72.8% reported one or more doctor visits within the past 12 months. A main strength of this study is that it used the NHIS to identify those who do not consult for their arthritis, and is a reliable source for self-report health data. The limitations were that data was not confirmed by a doctor and may be subject to recall bias. This study suggests that better health status may influence nonconsultation for arthritis. This finding is in agreement with Ethgen et al. (2002).

In the Ethgen et al. (2002) study, OA patients completed a minimum of 2 and a maximum of 6 surveys at 6-month intervals over a period of 2.5-3 years.

Longitudinal data analysis showed that health related quality of life (HRQOL) instruments (SF-36 (physical component score), the Stanford Health Assessment Questionnaire (HAQ) and the Western Ontario McMaster Universities Osteoarthritis Index (WOMAC) were all linked to future healthcare resource consumption. Respondents with lower HRQOL (i.e. worse health) scores had higher rates of healthcare resource utilisation compared to those in the higher quarters (i.e. better health).

Dominick et al. (2004) also found that greater HRQOL was associated with future healthcare use. Dominick et al. (2004) recruited a sample of older adults with OA in the U.S and mailed a version of the Centres for Disease Control and Prevention HRQOL modules (including recent days of 'not good' mental health/activity limitation/activity limiting pain/poor sleep and general health). Increased pain frequency was associated with visiting a physician (in the past 12 months). Pain frequency was the strongest predictor of all HCU variables among older adults with OA. The generalisability of results is limited due to an over representation of women and under representation of minorities, patients with less severe arthritis and users of self-care or complementary therapies, compared to the US population. Issues with coding accuracy meant a definite diagnosis of OA was unclear.

Hopman-Rock et al. (1997) described the pattern of HCU of patients aged 55-74 years with knee or hip pain who presented at different levels of the healthcare system in the Netherlands. Illness related variables and self-reported diagnoses were compared for attenders and non-attenders of GPs. Of the 82% who consulted a GP, 69% of them had arthritis and 65% of those attended a specialist. Those who were diagnosed with arthritis and had attended a specialist

were different from those who did not with regard to body mass index, chronicity of pain and attendance of a physiotherapist. No statistically significant differences were found in pain severity, level of disability, age or having a diagnosis of radiological OA. A limitation was that diagnoses were self-reported and it was unclear which criteria the doctor had used for diagnosis. A main strength was that a reference group of patients (patients attending GPs in other parts of the Netherlands for OA) was used, enhancing generalisability.

Rosemann et al. (2007), contrary to Hopman-Rock et al. (1997) found that increased pain severity was associated with HCU (i.e. a higher frequency of GP visits in the past 6 months). The study was a cross-sectional survey among patients with knee or hip OA from 75 primary care practices in Germany and had a response rate of 82%. Being single, a higher number of drug prescriptions, a higher score on the Patient Health Questionnaire (PHQ-9) (reflecting low mood or a depressive disorder) were also associated with a higher frequency of GP visits. The impact of psychological factors upon use of healthcare providers is a finding consistent with the De Boer et al. (1997) systematic review.

It has become evident from reviewing the six (non-UK) studies that there are some contradictory findings about the importance of a broad range of potential determinants for HCU. The contradictory findings may be attributable to the different ways in which these determinants have been measured. The factors associated with GP consultation (UK) for joint pain will now be presented in turn (see Table 1.2 overleaf), before a summary of these findings and implications is presented in the summary section (section 1.5.3 on page 21).

Factors associated with GP consultation (UK) for joint pain

Jordan et al. (2006) conducted a population-based prospective cohort study linking baseline self-report data to primary care medical records. The aim was to determine predictors of new episodes of consultation in primary care among older adults with knee pain. Respondents were aged 50 years or above, had reported knee pain in the past 12 months and had not consulted for knee pain in the last 18 months. Lack of social support increased the likelihood of future consultation – a finding that contradicts the De Boer et al (1997) review. Moreover, apart from pain chronicity, severity of pain and disability were not a strong influence on consultation (a finding consistent with Rosemann et al. 2007).

Table 1.2: Factors associated with GP consultation (UK) for joint pain

Andersen-Newman (1995) framework factors	GP consultation (UK)
Predisposing	Illness perceptions (Mitchell et al. 2006, Bedson et al. 2007, Hill et al. 2007) Social score (Mitchell et al. 2006) Urban living (Thorstensson et al. 2009)
Need-related	Depression (Jordan et al. 2006) Pain chronicity (Jordan et al. 2006) Previous injury (Jordan et al. 2006) Disease severity (Mitchell et al. 2006) Pain severity (Bedson et al. 2007, Thorstensson et al. 2009) Level of disability (Mitchell et al. 2006) Recent onset of pain (Bedson et al. 2007) Obesity (Thorstensson et al. 2009) Mobility problems (Thorstensson et al. 2009) Reporting three or more symptoms (Hill et al. 2007)
Enabling	Social support (Jordan et al. 2006) GP practice (Mitchell et al. 2006) Previous experience of healthcare (Jordan et al. 2006)

Current depression reduced the likelihood of consultation but the main predictor was previous experiences of healthcare (a finding consistent with De Boer et al. 1997 and Cronan et al, 1995). A main strength was the high response

rate (77%) and high consent to medical record review (79%). Generalisability may be limited by the patients living in one region of the UK (North Staffordshire).

Mitchell et al. (2006) conducted a cross-sectional postal survey including demographic, clinical (measures of pain, function, handicap) psychological (illness beliefs) and service use data to investigate the factors associated with GP consultation (in the past 12 months) in patients aged 50 and over with knee pain. Of the 231 patients reporting knee pain, 58% had seen their GP about it. Consulters had more severe disease, higher levels of (WOMAC) pain and disability and believed that their condition was more severe and had major consequences on their lives, compared to nonconsulters. A low response rate (34%) meant that the final sample size may not be large enough to assess associations between variables. Generalisability may be limited due to the sample being restricted to patients resident in one region of the UK. Although pain and function were important the strongest factors influencing GP consultation were GP practice, social score and illness beliefs. Thus enabling factors emerged as important in HCU, something which was not drawn out in the De Boer et al. (1997) review.

In agreement with Mitchell et al. (2006) Bedson et al. (2007) found that those who believed their knee problem was a priority were more likely to consult. Bedson et al. (2007) linked previous records of knee related consultations in general practice in the 18 months before baseline assessment of individuals aged 50 and above reporting knee pain in the past 12 months. Out of the 742 patients assessed, 28% had consulted for knee pain in the last 18 months. Based on odds ratios the odds of consulting for knee pain increased with recent onset of pain and severity of pain. However, although need factors were strongly associated with

consultation for knee pain, 50% of those with severely disabling knee pain did not consult for it. Irrespective of knee pain severity, self-reported comorbid conditions and the number of comorbid conditions were not associated with consultations for knee pain (a finding consistent with De Boer et al. 1997). Omitted factors that may have had more influence were previous HCU, feelings of personal risk. Multiple tests of association were used, thus increasing the risk of a type I error (i.e. the risk of obtaining a statistically significant result by chance). Thus, the findings might need to be considered in light of multiple tests of association.

In agreement with Mitchell et al. (2006) and Bedson et al. (2007), Hill et al. (2007) found that GP consultation was associated with patients' illness perceptions regarding their hand problems. These included the reporting of three or more symptoms, reporting more severe consequences, the belief that treatment could control their hand problems or pain, frustration and increased emotional representations. Hill et al. (2007) conducted a cross-sectional postal survey to investigate the illness perceptions associated with GP consultation (in the past 12 months) in adults aged 50 years and over with hand problems. A health survey questionnaire was sent to 11,230 patients from three general practices in North Staffordshire. Of the 7878 respondents to the health survey 47.6% (n=3749) reported hand problems. These patients were sent a hand questionnaire, including the Arthritis Impact Measurement Scales 2 (AIMS2), questions on self-reported diagnoses and GP consultation. A limitation with cross-sectional designs is that they cannot identify a direct causal link between illness perceptions and outcome.

Consistent with Mitchell et al. (2006) Bedson et al. (2007) and Hill et al. (2007), Jinks et al. (2007) found that beliefs about (knee) pain (i.e. it is a part of normal ageing), beliefs that few effective treatments are available and that

medication causes dependency, were key reasons for nonconsultation. Another reason for nonconsultation was that older people who lived with more than one condition ranked them in terms of urgency and severity. Pain intensity and perceived high impact on daily life were commonly cited reasons for consultation (a finding consistent with Mitchell et al. 2006). 53% of respondents with severe pain or disability had not consulted their GP in the last 12 months. Needs assessment was defined in terms of 'felt' need (individual assessment of a need for healthcare) and 'expressed' need (demand for healthcare). A population survey including a Knee Pain Screening Tool, HCU and the Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) was sent to adults aged 50 and above from three general practices. Based on the WOMAC scale respondents were classified as 'severe' or 'extreme'. A strength of this study was that beyond the epidemiological data qualitative interviews and patient diaries were used to detail the experience of living with knee pain, disability and consultation with a GP.

Finally, Thorstensson et al. (2009) found that disability was the strongest determinant of GP consultations among adults aged 35 or over living with chronic hip or knee pain in a large community sample. This finding contradicts the finding by Jordan et al. (2006) that disability was not a strong determinant of consultation. However, results also suggested that mobility problems, urban living, obesity and pain severity were associated with HCU – a finding (i.e. the importance of pain severity) that was consistent with Bedson et al. (2007). Thorstensson et al. (2009) defined HCU by visiting the GP in the previous 12 months or an allied health professional (AHP) in the previous 3 months. Of those reporting hip or knee pain, 52% had pain at both sites. 25% only consulted a GP, 3% an AHP only, and 4% an alternative therapist only. 13 % consulted more than one of these categories,

and 55% had not seen any health professional. The study response rate was high (86%) and a large, representative sample was recruited. However, the data on help-seeking behaviours and comorbidities was self-reported which may imply recall bias (i.e. there may be systematic error due to differences in accuracy or completeness of recall to memory of previous help-seeking or comorbidities). For example, the information recalled by respondents is more likely to be reliable about alternative practitioners (in the last 3 months) than those about GPs (in the last 12 months).

1.5.3 Summary

The overview of empirical evidence to date suggests that:

- there are contradictory findings about the importance of determinants for GP consultation.
- there are a broad range of potential determinants, indicating that the decision-making process is complex.
- the contradictory findings may be attributable to the different ways in which these determinants were measured.

It may be that the designs used to date do not capture the complex decision-making process itself. These findings will be discussed in the following section.

It is clear that there are contradictory findings about the importance of determinants for GP consultation. For example, in one study disability was not found to be a strong predictor of consultation (Jordan et al. 2006), and in another it was a large predictor of consultation (Thorstensson et al. 2009). The early De Boer et al. (1997) review suggested that social factors were not significant determinants of consultation, though social support was a significant predictor of consultation in a later study (Jordan et al. 2006).

Perhaps most interesting was the mixed evidence surrounding the influence of pain severity upon consultation. Although pain is without doubt the most salient presenting feature of symptomatic OA (Loeser 2000), it appears that evidence from quantitative observational studies suggest that painful symptoms alone are insufficient in the decision to consult a healthcare professional (Jordan et al. 2006, Thorstensson et al. 2009).

Instead, it appears as though many people with longstanding joint pain may become resigned to it, attributing their problems to 'normal ageing' and having low expectations and negative experiences of primary care (Jinks et al. (2007). This is something reflected in wider literature (Stoller et al. 1993, Sanders et al. 2004). Another reason for nonconsultation suggested was that many older people who lived with more than one condition ranked them in terms of severity and perceived urgency (Jinks et al. 2007). However, in other studies the evidence for comorbidities pushing (knee) problems down the agenda for consultation was unclear (Bedson et al. 2007, De Boer et al. 1997).

There was some interesting and consistent evidence about the importance of illness perceptions (Mitchell et al. 2006, Bedson et al. 2007, Hill et al. 2007, Jinks et al. 2007). For example**,** whether respondents perceived their knee problem as an important health issue (Bedson et al. 2007), was most strongly associated with consultation. Jordan et al. (2006) and Cronan et al. (1995) found that previous experience of the health service is a consistent predictor of subsequent consultation. De Boer et al. (1997), despite suggesting that enabling factors have little impact upon HCU, concluded that in order to understand HCU behaviour an understanding of the impact of provider and the healthcare environment upon individuals is needed (de Boer et al. 1997). This is something that has been

relatively under researched using this model. Few of these studies included important determinants of consultation such as the attitude and expectations of patients towards the technical and interpersonal skills of the care provider (Bradley et al. 2002)

From the current empirical evidence it is clear that there is a broad range of potential determinants. This indicates that due to the ways in which patients perceive, evaluate and manage their symptoms, the decision-making process for OA consultation is complex (Campbell & Roland 1996). It clearly involves the process of patients weighing up numerous factors. Although the prevailing biopsychosocial Andersen-Newman framework is comprehensive, it does not fully explore how these domains or measures interrelate in a dynamic internal process Arguably, the complexity of the decision-making process is difficult to capture in research studies (particularly quantitatively).

Given the limitations of observational epidemiological studies for understanding the decision-making process, another approach to understanding GP consultation has been to use qualitative research. For example, two of the reviewed studies suggested that the strongest predictors of GP consultation were previous experiences of healthcare (Cronan et al. 1995, Jordan et al. 2006). This is likely to be a reflection of patients' expectations or experience of the healthcare system and healthcare providers. This finding can perhaps be illuminated by findings from qualitative research into OA consultation. For example, Gignac et al. (2006) found that for many OA patients OA symptoms were perceived as *"an inevitable part of the ageing process requiring acceptance and not treatment"*– a belief that was influenced by health professionals' pronouncements that nothing could be done for OA (Gignac et al. 2006, p.1). In the wider general practice and

consultation literature the impact of patient-centred GP attitudes and the GPs' interpersonal skills (listening and communication skills) upon the patients' experience of the consultation is nothing new (Grol et al. 1990). The GP attitude is defined as taking the patient's ideas and expectations seriously and giving information to enable the patient to take responsibility for their own health (Grol et al. 1990). Although some of these issues have not been investigated in relation to OA consultation extensively, in OA management research qualitative research is emerging which includes reports by patients that the GP does not take their OA seriously (Alami et al. 2011). Patients believe that knee OA is an inevitable part of ageing and, that little can be done by practitioners (in terms of treatment) to modify its development (Alami et al. 2011). The latest OA guidelines also show that OA patients have reported barriers to communicating with doctors and that some feel that their health professionals ignore their symptoms (Gignac et al. 2006). This might reinforce beliefs that OA is an inevitable and normal part of ageing and that health professionals might consider them a burden on the NHS (National collaborating centre for chronic conditions 2008).

Overall, there are a number of complex factors involved in the decision to consult the GP for joint pain and it appears that decision-making is a process difficult to capture using observational epidemiological designs. It is clear that quantitative designs provide comprehensive lists of determinants with little understanding or explanation of the dynamic process of decision-making itself and the importance of these factors within that process. Moreover, although qualitative studies reveal the importance of complex factors of service, they are not able to give quantitative estimates of the relative importance of them. The next section will

consider the strengths and weaknesses of the study designs used in generating the empirical evidence for determinants of consultation to date.

1.6 Methodological limitations and gaps in knowledge

The principal designs used to investigate determinants of consultation for joint pain/OA have been cross-sectional surveys or longitudinal (prospective or retrospective) observational studies. These studies essentially model observed behaviour. This has the major strength of dealing with actual consultation behaviour in people 'as they are' rather than hypothetical scenarios. However, there are limitations. Service and service provider characteristics are often not included as determinants as these are effectively fixed: the services and GPs are as they are. There is little scope therefore for investigating whether or not consultation behaviour would change if service characteristics were changed (e.g. if more effective treatments were available). A similar issue arises when considering patient characteristics that vary over time (e.g. pain intensity, the severity and intrusiveness of comorbid illness). In traditional quantitative observational studies, each person contributes only one observation - the combination of their attributes at that point in time and those of the service(s) they have access to. These designs therefore rely on between-person differences. It may require a very large study involving many different types of patients and practices to be able to evaluate the relative importance of such a wide range of factors. Furthermore, it can be seen that observational studies are poorly adapted for more hypothetical 'what if?' questions, particularly in relation to changes in service configuration or provision. These include combinations of services (that may not currently be available) that will encourage those most in 'need' of care to consult the GP for joint pain. To answer questions such as - whether under altered

conditions the same person would be more or less likely to consult? – require other research designs.

Alternative methods exist that may address these above limitations. Utility-based stated preference (SP) (rather than revealed preferences: i.e. what people actually do) methods such as conjoint analysis are now well validated in health research (Stiggelbout et al. 2008). These methods have been used to uncover preferences across a variety of contexts, including treatment preferences for osteoarthritis (Fraenkel et al. 2004a), to aspects of patient-centred care in primary care consultatons (Haas 2005, Cheraghi-Sohi et al. 2008).

The next chapter introduces conjoint analysis (CA) and outlines its theoretical history, validity in health research and its potential for explaining the determinants of GP consultation for osteoarthritis.

2 Chapter Two: Conjoint analysis as an approach for investigating the determinants of GP consultation

2.1 Synopsis

Chapter one considered the findings and methodological limitations of previous research on the determinants of GP consultation for symptomatic osteoarthritis. The first section of this chapter introduces conjoint analysis techniques and gives an overview of these methods and their validity in marketing and health research (section 2.2). The next section gives an overview of the advantages and disadvantages of conjoint analysis in health research (section 2.3), before outlining the reasons for choosing conjoint analysis to understand the decision to consult the GP for joint pain (section 2.4). Finally, the thesis aims and objectives are stated (section 2.5) before the structure of the thesis is presented and detailed (section 2.6).

2.2 Conjoint analysis: what is it?

Conjoint analysis (CA) is a general term that has come to encompass a range of quantitative methods of eliciting preferences (Ryan 1999, Ryan et al. 2003). Conjoint analysis methods (and discrete choice experiments, DCEs) are widely used stated preference (SP) elicitation methods across disciplines such as marketing, environmental and health economics (Louviere et al. 2000). Conjoint analysis is based on the simple premise that people evaluate the overall utility (value or desirability) of a product or service by combining the utility provided by each attribute characterising it. It works by asking respondents to rank, rate or choose between hypothetical products or services that are presented to them (Ryan et al. 2001). These hypothetical evaluations are then used to infer the utilities of the individual attributes comprising the product or service. CA allows a

mathematical estimation of the relative importance of different attributes of a product or service, the trade-offs between these attributes presented at different levels and the total satisfaction or utility that respondents derive from a product or service (Ryan et al. 2001, Ryan & Farrar 2000). Conjoint methods are considered preferable to traditional methods because instead of asking respondents directly to rate how well a service performs on a number of attributes, they (particularly choice-based conjoint) force respondents to make a choice - thus getting at their underlying priorities more precisely. Conjoint also reduces the chances of obtaining a socially desirable response (i.e. respondents providing answers that they think the researcher wants to hear) (Flach & Diener 2004).

CA evolved from a method called conjoint measurement, which was originally developed in mathematical psychology by Luce and Tukey (1964) with the aim of quantifying human preferences (Luce & Tukey 1964). According to Green and Rao, (1971) conjoint measurement is "*concerned with the joint effect of two or more independent variables on the ordering of a dependent variable'* in a marketing context (Green & Rao 1971, p. 355). Conversely, discrete choice experiments (DCEs) came from econometrics, building upon the work of McFadden (1974), who won the Nobel Prize in Economics in 2000 for his work in the 1970s and 80s on choice modelling (McFadden 1974). Both SP methods, CA and DCEs (often used synonymously) have gradually been applied increasingly to other contexts, notably healthcare.

2.2.1 Conjoint analysis: how does it work?

Conjoint analysis is based on three fundamental assumptions:

- Each product or service is a bundle of potential attributes.

- Each individual has a set of unique internal relative utility weights for attribute levels.
- Combining the utilities for different attributes provides an individual's overall relative utility (Singh et al. 1998).

The different CA formats will be explained using the attributes and levels from a published study investigating preferences for primary care consultations (Caldow et al. 2007). Firstly, in CA each product or service is (made up of) a bundle of potential attributes (also called factors or determinants). For example, a patient may evaluate the following primary care consultation service alternative/profile in Table 2.1. This is made up of the following service attributes: who you see, waiting time for appointment, length of consultation, continuity of health professional and likelihood of having illness cured (Caldow et al. 2007) (see Table 2.1). Each attribute has several possible levels. In this example, the level of 'who you see' is practice nurse.

Table 2.1: Worked example of possible attribute levels for primary care consultation alternative

Attribute	Who you see	Waiting time for appointment	Length of consultation	Continuity of health professional	Likelihood of having illness cured
Levels	Practice nurse	4 days	20 min	Yes	75%

Secondly, each individual holds a set of unique preferences (partworth utility values) for attribute levels. CA assumes that each individual has an internal set of values (that follow the rule of additivity) to evaluate the overall desirability of a product or service. The rule of additivity suggests that respondents *'apply a compensatory rule when considering which product to buy/choose, wherein bad*

characteristics can be overcome by the presence of other very good characteristics' (Orme 2006, p.104). This simplistic view of behaviour is criticised by Louviere et al. (2011), who views this as imposing restrictions on one's ability to understand the true nature of decision and valuation processes (Louviere et al. 2010). An example of a given patient's internal additive values (or partworth utilities) for a health service (a primary care consultation) is shown in Table 2.2. In Table 2.2 the continuity of the health professional is the most important feature for this respondent, with 50 internal partworth utility points.

Table 2.2: Worked example of a respondents internal partworth weights

Attribute	Who you see	Waiting time for appointm-ent	Length of consultati-on	Continuity of health professio-nal	Likelihood of having illness cured	Total utility
Levels	Practice nurse	4 days	20 min	Yes	75%	
Partworth	20	5	15	50	30	120

Finally, combining the value of the different attributes provides an individuals' overall relative preference. In other words CA assumes that the overall desirability of a health service is equal to the sum of its parts (Orme 2006). The function of CA for product or service marketing is to obtain these internal partworth weights for a variety of attributes that may be used in a product or service design.

The advantage of CA is that the researcher can select the salient attributes of interest and input them into a CA design that will systematically vary the features of the service alternatives. They can present the task to respondents for evaluation and then statistically deduce (using multiple regression methods) the utility scores (partworth utilities) for each of the attributes in order to understand what rules respondents unconsciously use to evaluate the services (Lancsar &

Louviere 2008). These partworth utilities can then be used to measure the desirability of a service (like a primary care consultation) and predict how respondents would choose between new configurations of services.

The design and analysis of conjoint analysis/ discrete choice experiments have been defined as consisting of five steps (Ryan & Gerard 2003) (see Table 2.3 below).

Table 2.3: Stages of Conjoint Analysis/DCE

Stage	Description
1.	Selection of salient attributes, levels, and scenarios
2.	Experimental design and construction of choice sets
3.	Measurement of preferences
4.	Estimation procedure
5.	Tests of the validity of responses

These stages will now be described using the same primary care consultation example (Caldow et al. 2007) as in the previous section.

Selection of salient attributes, levels, and scenarios

To obtain CA data (i.e. respondents internal values), a list of salient attributes are defined, typically based on literature reviews and qualitative research like focus groups, experts reviews and pilot testing (Bridges et al. 2011). Table 2.4 below shows a list of hypothetical attributes salient to patient preferences for a primary care consultation.

Table 2.4: Example of a list of hypothetical attributes and levels

Attribute	Who you see	Waiting time to appointment	Length of consultation	Continuity of health professional	Likelihood of having illness cured
Levels	Doctor Practice nurse	No waiting time 2 days 4 days 8 days	5 min 10 min 20 min 30 min	Yes No	75% 80% 85%

Attribute levels represent relevant possibilities given existing and potential future developments in primary care consultations and must be identified and reviewed for relevance and comprehension by experts and focus groups. The scenario often refers to the preference task that respondents are being asked about and how it is framed and worded, e.g. *'For the past three or four mornings you have coughed up a little phlegm and you decide to ask for an appointment at your practice'* (Caldow et al. 2007, p.32). Although, in other studies (and this study) the term 'scenario' refers to the individual service alternative/profile being presented in the choice tasks (See Figure 2.1 on page 33 for an example of a service alternative).

Experimental design and construction of choice sets

Once the salient attributes are selected (as in Table 2.4) these are used to create hypothetical primary care consultation alternatives/profiles (sometimes called scenarios). It is clear from Table 2.4 that there are many possible combinations that can be generated from these attribute levels (2*4*4*2*3) = 192 possible alternatives. However, it is too cognitively demanding to ask respondents to evaluate all of these 192 alternatives (a full-factorial design). Instead, assuming an additive model, respondents only need to evaluate a reduced amount of the total 192 combinations (a fractional-factorial design). Fractional-factorial designs are constructed from a fraction of the total combination and based on orthogonal profiles. Orthogonality within CA refers to experimental designs in which the attribute levels are uncorrelated across consultation alternatives (Bridges et al. 2011). These designs are based on an algorithm and can be generated using published designs or statistical software like Sawtooth Software (Inc. Orem, UT) (Bridges et al. 2011).

Bridges et al. (2011) recommended potential criteria for evaluating statistically efficient designs:

- Efficiency scores (how precisely a given set of conjoint questions can estimate the parameters of interest). See chapter 5 for more information.
- Attribute levels are uncorrelated across consultation alternatives (orthogonality).
- Level balance (attribute levels occurs an equal number of times).
- Minimal level overlap (attribute levels appear once in a choice alternative).
- Restriction of implausible combinations of levels in a choice task (i.e. no attribute levels should directly contradict other attribute levels).
- Cognitive burden of respondents should be minimised (Bridges et al. 2011).

In this example in order to calculate a full set of partworth utilities for each respondent for the 5 attributes covering 15 levels the guidelines recommend showing 1.5 to 3 times as many consultation alternatives as you have parameters to estimate (number of parameters = total number of levels - number of attributes + 1, in this case = 11 parameters). Thus respondents are required to evaluate a minimum of 16-33 consultation alternatives. Figure 2.1 shows an example of one of these consultation alternatives (Orme 2006).

Figure 2.1: Example of a primary care consultation alternative in the rating format

How likely are you to prefer this primary care consultation? Use a scale from 0 – 100: Where 0= not at all likely and 100 = definitely would prefer

- **With a practice nurse**
- **4 days waiting time for appointment**
- **20 mins long**
- **Continuity of health professional - Yes**
- **75% likelihood of having illness cured**

Measurement of preferences

Service or product alternatives (like the primary care consultation example) can be evaluated in three different formats: ranking, rating, and choice. During the infancy of conjoint analysis it was popular to ask respondents to evaluate many product alternatives printed on cards by ranking (or rating) them. The rating format is shown in Figure 2.1. However, currently choice-based CA is most commonly used as opposed to rating or ranking (Orme 2006, Marshall et al. 2010). This is because making choices is an easy and familiar real-world choice behaviour. In choice-based conjoint (CBC) (see Figure 2.2 on page 42) orthogonal arrays are used as seed profiles and the choice alternatives for CBC surveys are generated from the seed design.

Instead of allowing respondents to directly evaluate each attribute level as equal, as is possible in traditional rating exercises, CBC forces individuals to trade-off the different attributes of the service, weighing up services that have more and less desirable characteristics (Phillips et al. 2002a). Therefore, when respondents are forced to make difficult trade-offs, researchers are able to learn about their precise values and preferences.

Estimation procedure

The partworth utilities of these 16-33 primary care consultation alternatives are then evaluated by respondents and these evaluations are used to understand which levels are preferred and the relative importance of each attribute. In rating card sort exercises the utility estimate for each level is the average score for cards that include that level. However, in CBC survey responses are interpreted as utility differences between the choices. In CBC, *'a dataset is thus created where the dependent variable is the person's choice and the independent variables are the*

difference in the attribute levels in the choice that the respondent saw' (Phillips et al. 2002b*,* p.1700). Probit or logit analysis is typically used to fit probabilities to the choices made by respondents (Phillips et al. 2002b).

The relative importance of each attribute is calculated by dividing the range of utilities for each attribute by the sum of the ranges and multiplying by 100 (Orme 2006). Once the partworth utilities are added together to determine the relative importance scores, it is possible to add them together to predict how each respondent will respond to the remaining combinations (Orme 2006). CBC utilities can also be computed for each level of each attribute for each respondent, using Hierarchical Bayes modeling (Kievit et al. 2010). See section 6.4.2 on page 182 for a full explanation of CBC data analysis methods.

Tests of the validity of responses

In CBC, validity of estimates is addressed in two ways. Firstly, holdout tasks are used to evaluate internal validity, i.e. whether utilities estimated from scenarios can accurately predict some additional choice or conjoint question not used in the estimation of partworth utilities. Holdouts look identical to standard CBC questions, typically 3 to 5 service alternatives, though these are not used to estimate partworth utilities. These are 'held out' in order to assess the performance of the estimated partworths. The model can be validated when the responses to holdout questions are accurately predicted using estimated partworths (Orme et al. 1997).

Secondly, observational data (or revealed preference data) is used to ascertain whether the CA utilities predict actual behaviour (external validity). External validity refers to the model's ability to predict events outside of the survey, for example actual market purchases (Orme 2006).

The amount and pattern of missing data (i.e. data is missing when a respondent has not indicated a choice on the questionnaire) may also provide some insight into respondent burden. In discrete choice experiments tests of rationality are also used to assess the validity of utility estimates. Based on the assumption that individuals attempt to maximise the utility they obtain and are willing to trade between choices in order to maximise, rationality is evaluated using dominant options. A dominant option is a choice alternative which contains attribute levels that are superior on all attribute levels to a random comparison alternative. If respondents do not maximise and select a dominant alternative then their choices are considered irrational (Ryan et al. 2009).

Reliability refers to the consistency of respondents' evaluative processes and may be evaluated using repeated choice tasks (i.e. the same choice task is presented at the beginning and end of the questionnaire to check for consistency). However, a higher reliability does not necessarily lead to more precise models because respondents may complete a conjoint study reliably but may not make realistic choices, or the conjoint questionnaire may not contain the salient attributes. For example, respondents who use simplifying strategies (e.g. always picking the same attributes) have higher reliability scores than those who calmly weigh up all attributes before choosing or rating a product/service (Orme 2006).

Furthermore, there are studies that compare the estimates from conjoint estimates with traditional (Likert) rating and ranking scales. Some studies suggest that there is discordance between simple rating and choice-based conjoint methods (Hundley & Ryan 2004, Bridges et al. 2010). This is explained by the way in which in conjoint studies respondents are typically asked to make choices within

a resource constraint, but when evaluating Likert scales, respondents are able to rate all attributes of a service or product as equal. CBC thus forces individuals to trade-off more and less desirable characteristics of services (Phillips et al. 2002a). However, other studies suggest that the relative importance of attributes from DCE results were entirely consistent with the order of attributes from a simple ranking exercise (Peacock et al. 2006). Phillips et al. (2002a) found simple ranking and rating scales are generally consistent with conjoint ranking and rating scales respectively. However, in the Phillips et al. (2002a) study focus groups revealed that respondents believed that the conjoint analysis tasks were useful at forcing them to think more deeply, and they reported that the conjoint results more accurately reflect how they would behave in the real world (Phillips et al. 2002a).

2.2.2 Theoretical development of conjoint analysis in marketing and healthcare research

Those most familiar with the development of CA, including Orme (2006) – providers of statistical Sawtooth Software (Inc. Orem, UT) - argue that CA sprung out of a need in marketing research in the 1970s to predict (rather than describe) what consumers will buy when faced with potential new products with a range of different attributes (Green & Rao 1971, Orme 2006). Consequently, Professor Paul Green developed conjoint measurement (Luce & Tukey 1964) into CA and as CA began to expand in marketing in the early 70s, many innovations of conjoint measurement started to emerge. Innovations included reducing full-factorial designs to fractional-factorial (reduced designs) to elicit the relative importance of attribute levels; using rating scales instead of rankings to elicit preference orders and using statistical models to estimate preference parameters (i.e. partworth utilities) (Louviere et al. 2010).

CA has been widely applied in economics and marketing research (Hensher et al. 2005), although extending these methods to abstract concepts such as treatments and services in healthcare (rather than more familiar products used in marketing research) may not be achievable (Johnson 2008). Conjoint analysis was first utilised in a healthcare setting (particularly with the use of discrete choice experiments in health economics) as a technique to extend beyond the quality adjusted life year (QALY) approach, a method that quantifies health outcomes (Ryan et al. 2005). Beyond health outcomes researchers believed that other 'service' based aspects of healthcare were important for policy decisions, for example, waiting time (10, 20 or 30 minutes) and type of healthcare professional (Ryan et al. 2005).

2.2.3 Formats of conjoint analysis used in marketing and healthcare research

Since Green's early adaptation of conjoint measurement (1974), CA has become a widely used method and a range of adaptations are currently in use. These include traditional full-profile conjoint analysis, i.e. conjoint value analysis (CVA), adaptive conjoint analysis (ACA) choice-based conjoint analysis (CBC) – including partial-profile choice-based conjoint (PPCBC). According to Sawtooth Software (Inc. Orem, UT), the dominant conjoint techniques are full-profile CVA (10%), adaptive conjoint analysis (20%) and choice-based conjoint (70%) (Orme 2006). Almost 80% of the studies reviewed in 2008 were choice-based (Marshall et al. 2010). Only the primary techniques will be considered in this thesis, including the recently developed adaptive choice-based conjoint (ACBC).

Rating (or ranking)-based CA (CVA and ACA)

Rating (or ranking) based CA has been the leading conjoint technique since CA originated in the 1970's. In rating techniques respondents rate each attribute level on a scale where the upper measurement is 'definitely would prefer' (a score of 100) and the lower measurement is 'definitely would not prefer' (score of 0) (see Figure 2.1 on page 33). However, in ranking techniques individuals are asked to rank individual scenarios involving different combinations of levels in order from best (1) to worst (10). For example, Rosko et al. (1983) asked 97 university students to rank in order of preference 26 service alternatives (written on cards) which included ambulatory health service attributes including: charge for routine visit, office hours, length of time needed to make an appointment (Rosko et al. 1983).

Full-profile CA (CVA) appears in two different formats: single profile (ranking or rating), and pairwise comparison (rating only). A pairwise comparison format presents two product alternatives and the respondent is able to indicate a strength of preference on a rating scale ranging from 2 to 9 points (Orme 2006). CVA is presented as a full-profile, which means that each product alternative/profile includes all of the attribute levels being studied. CVA thus does not include more than six attributes in each profile to avoid cognitive burden (Bridges et al. 2011, Pearmain et al. 1991). It can be computer or pen-and-paper based.

Ranking-based CA has been applied in various health contexts from alternative contraceptive methods (Nickerson et al. 1991) to preferences for growth augmentation therapy (Singh et al. 1998). Singh et al. (1998) demonstrate that respondents did not find any difficulty completing ranking exercises and

another study into rural primary healthcare facilities took respondents up to 18 minutes to rank 25 alternatives (Parker & Srinivasan 1976). Ranking CA has been least popular in healthcare (Ryan et al. 2001).

Adaptive Conjoint Analysis (ACA) was the first computer-based method and was most popular in the 1990's (Orme 2006). ACA tasks involve three question types. Firstly, respondents rank the levels within each attribute. Secondly, respondents rate the importance of the difference between the best and worst of each characteristic on a 4-point scale (1, not important at all; 2, somewhat important; 3, very important; 4, extremely important). For example: *'If two products were acceptable in all other ways, how important would this difference be?: 'Starts to work within 1 to 2 hours' versus 'starts to work within 4 weeks'* (Orme 2006, p. 35). Finally, respondents evaluate a series of pairwise comparisons on scale of 1 to 9, where 1 indicates strongly prefer; 5, no preference; and 9, strongly prefer.

ACA has been applied to HIV treatment preferences within differing medication contexts (Beusterien et al. 2005), assessing preferences over time and capturing individual preferences as distinct from treatment experience in cancer patients (Pieterse et al. 2010) and treatment preferences for knee OA (Fraenkel et al. 2004a). However, evaluation of its validity and reliability is limited.

ACA is able to customise the conjoint task via a computer algorithm that allocates choices based on previous responses. This allows researchers to get a better understanding of the most important attribute levels for individual respondents (Orme 2006). ACA is also able to include more attributes than CVA or CBC. The guideline is that respondents cannot efficiently process more than six attributes at one time in a full-profile task (Pearmain et al. 1991). However, there is evidence to suggest that this is influenced by additional factors like length of

attribute as well as familiarity of the attributes to the respondent (Orme 2006). Sawtooth software claims that ACA can include 30 attributes, though ACA studies typically range from 8 to 15 (Orme 2006). One limitation of ACA is that it does not imitate choice behaviour, i.e. respondents rarely examine 10 to 20 alternatives and rank them individually (Stiggelbout et al. 2008) in the real world. However, the main disadvantage of ACA is that in order for it to adapt to individuals preferences it must be administered by computer. This could be a problem for older age groups where computer access and computer literacy may be limited.

Choice-Based Conjoint (CBC and ACBC)

Based on the methods used in published works by practitioners and academics it seems that since the 90's choice-based conjoint has been used more frequently than rating (or ranking) based methods. In CBC individuals are presented with choices that involve different combinations of a product/service, and are asked to indicate which one they would choose (see Figure 2.2 taken from Caldow et al. 2007). The response options are either in a choice format (i.e. prefer A or B) or graded (i.e. strongly prefer A, prefer A, indifferent, prefer B, strongly prefer B). CBC designs can also offer respondents a choice to 'opt-out' and not make a choice as is possible in the real world marketplace (as with option C in Figure 2.2 overleaf).

Figure 2.2: Example of primary care consultation alternatives in the discrete choice format

If you were offered options A, B and C below, which one would you choose?

A	B	C
• Consultation is with a **practice nurse** • **4 days** waiting time for appointment • **20 mins** long • Continuity of health professional -**Yes** • **75%** likelihood of having illness cured	• Consultation is with a **doctor** • **6 days** waiting time for appointment • **10 mins** long • Continuity of health professional -**No** • **85%** likelihood of having illness cured	• **None**: if these were my only options I'd defer my choice

Choice-based conjoint analysis (CBC), also commonly referred to as discrete choice experiments (DCEs), is popular because it imitates real-world choice behaviour. It simulates the psychological process more accurately because it is less abstract than ranking and rating (Huber 2005). In the same way that consumers do not individually rate or rank product/service alternatives when making a decision respondents (as with the real world) are forced to make choices. CBC can be pen-and-paper based or administered (online) by computer.

One approach to CBC is a partial-profile design (PPCBC). The aim of PPCBC designs is to increase the number of attributes that can be managed by respondents (Orme 2006). In PPCBC choice alternatives include a subset of the total number of attributes that are included in the design. '*These attributes are randomly rotated into the tasks, so across all tasks in the survey each respondent typically considers all attributes and levels'* (Orme 2006, p.40). The limitation of PPCBC is that the data is spread thinly because each choice task has attribute levels omitted. Therefore, although PPCBC enhances respondent efficiency it

reduces statistical efficiency because the individual level parameter estimates are less stable (Patterson & Chrzan 2004). Thus PPCBC requires larger sample sizes to stabilise results. There is also an assumption of 'ceteris paribus', i.e. researchers assume that respondents will hold equal all other attributes not included in the partial-profile choice sets. When deciding to use PPCBC, researchers must be willing to accept larger standard errors if it means greater respondent efficiency (Cunningham et al. 2008). Cunningham et al. (2008) used PPCBC to model the information preferences of parents seeking mental health services for 6 - 18 year olds. Using a PPCBC format enabled a total of 20 four-level attributes to be presented across 30 choice tasks without overburdening respondents. PPCBC has also been applied to developing patient-centred care health services (Cunningham et al. 2008) and medical education (Cunningham et al. 2006).

The most recent CBC method is Adaptive Choice-Based Conjoint (ACBC). ACBC claims to have the benefits of ACA and CBC because it is choice-based, able to manage more than 5 attributes, computerised and adapts to respondents' answers (Orme 2010). Despite requiring more time to complete than regular CBC or partial-profile CBC (Chapman et al. 2009), respondents consider ACBC methods as more engaging (Cunningham et al. 2010). Cunningham et al. (2010) recommend ACBC to assess health service preferences and reported that ACBC surveys produce lower standard errors and improved prediction of holdouts. Sawtooth Software (Inc. Orem, UT) research also suggests that ACBC gives more accurate predictions than CBC (Orme 2010). However, it is important to treat this early research with caution, especially considering the commercial interest that Sawtooth have in promoting the validity of ACBC.

2.2.4 Validity of conjoint analysis in marketing and healthcare research

Overall, CA methods are well validated in marketing research. For example, research into internal validity compares rating-based and choice-based conjoint approaches, using their ability to predict both aggregate choice shares among the sample and individual choices (Moore et al. 1998). There is also evidence in transportation economics (Adamowicz et al. 1994) that suggested a strong relationship between what people state they will do (stated preferences) and what they do in reality (revealed preferences). ACA (rating or ranking) is considered less realistic in terms of real-world buying behaviour (Orme 2006).

It is important to note that much of the reference material for CA is based on the work done by the providers of Sawtooth Software (Inc. Orem, UT). Similarly, other validation research is potentially industry sponsored. For example, Microsoft research into consumer electronic purchasing behaviour suggests that ACBC can predict real-world purchases more precisely than CBC (Chapman et al. 2009). Due to the sources of this research there is, to some degree, a question of trustworthiness. For example, systematic reviews suggest that industry sponsored trials have larger effect sizes than those done independently (Gluud 2006). Most recently ACBC has been recommended to assess health service preferences with reports that ACBC surveys produce lower standard errors and improved prediction of holdout tasks (Cunningham et al. 2010), though further validation from independent sources is needed.

In terms of independent research there are some fundamental criticisms that can be balanced against the vested interests of Sawtooth Software (Inc. Orem, UT). For example, different design features and data collection procedures may reduce the reliability and validity of CA methods, arguing that a full-profile

method leads to superior results (Darmon & Rouzies 1994). Perhaps the biggest criticism is that it is unclear whether what individual's state in a hypothetical context accurately represents their real-life decision-making (Kievit et al. 2010).

Healthcare research indicates a high level of respondent consistency and validity in CA rating exercises (Ryan et al. 2001). Completion rates range from 78 to 85% (Reardon & Pathak 1990, Harwood et al. 1994), with postal response rates ranging from 42 to 67% (Ryan et al. 2001). CA rating exercises have proven to lead to fatigue (Graf et al. 1993). Conversely, there is evidence to suggest that 16 to 24 alternatives/profiles can take up to 20 minutes to complete and that CA rating via face-to-face interviews are preferred to telephone interviews (Graf et al. 1993).

Ranking CA is also well validated in health research, achieving a 59% response rate in one postal survey (Carroll & Gagnon 1984). Respondent consistency and internal validity appear to be high (Ryan et al. 2001) though further research is needed. Few comparisons between the different CA methods have been made, with one study suggesting that rating and choice-based conjoint predict holdout tasks equally well but CBC simulates reality more effectively, leading to greater external validity (Elrod et al. 1992). It has been suggested that CBC methods are more difficult to complete than ranking because they demand a direct commitment of one's values (McClain & Rao 1974).

Choice-based CA has received response rates ranging from 18% to 88% (Ryan et al. 2001). Response rates are improved when conducted alongside trials, are clinic-based, offer monetary incentives or when a cover letter is sent from a doctor. It is recommended that no more than five or six attributes be presented at one time and no more than 12 choice sets in total (Ryan et al. 2001). Internal

validity has been high (Ryan et al. 2001), though one study showed evidence of an ordering effect (Vick & Scott 1998). Tests of rationality based on dominant options suggest that a small minority of people give irrational responses (Ryan et al.1999). San Miguel et al. (2000) investigated test–retest reliability of choice-based CA within a two month period in parent preferences for out-of-hours care and found high levels of reproducibility (San Miguel 2000).

2.3 Potential contribution of conjoint analysis

2.3.1 Advantages of using conjoint analysis in healthcare research

Potential advantages of using CA to understand determinants of GP consultation in OA include that it provides the relative importance of the determinants of GP consultation. This information can be used to develop new service configurations (Lancsar & Louviere 2008). While social desirability biases can affect results of health survey questions (Tourangeau & Yan 2007) in CA covert preferences can be accessed (Caruso et al. 2009).

CA uses a 'decompositional' approach, whereby respondents evaluate scenarios composed of a combination of attribute levels, with each level explicitly stated (e.g. attribute levels for waiting time until appointment; No waiting time, 2 days, 4 days, 8 days). Thus, the method allows utility to be estimated for each attribute level and for all possible combinations of levels, including combinations not directly evaluated by respondents (Phillips et al. 2002a). By contrast, individual attributes are evaluated as a whole (e.g. evaluating the importance of waiting time until appointment would be evaluated as a whole rather than separate levels on the importance of a 2 days, 4 days, 8 days waiting time for appointment) (Phillips et al. 2002a).

Additional advantages are summarised below in greater detail based on the work of Haider and Ewing (1990).

- CA allows for the design of experiments in which a large number of salient attributes can be combined to describe a hypothetical scenario, (in this case the decision to consult the GP) and in which research subjects evaluate (the decision to consult as a whole) instead of rating single attributes.
- CA allows researchers to control the alternatives and choice sets presented to respondents.
- CA allows different alternatives, including service provisions which may not currently exist. These can be designed and presented to respondents for evaluation (Haider & Ewing 1990).

Overall, research suggests that respondents are able to complete CA methods effectively (Viney et al. 2002). The potential of CA as an instrument for establishing the preferences of patients in the community has been recognised and application in the NHS recommended (Ryan & Farrar 2000). Essentially, CA methods use the findings from qualitative studies to adequately define the range of attributes that are relevant in the choice tasks. In this way CA studies complement other traditional approaches but unlike other methods they force respondents to make trade-offs, thereby getting respondents to think more deeply about their preferences (Phillips et al. 2002a).

2.3.2 Disadvantages of using conjoint analysis in healthcare research

The advantages of CA, however, must be considered by contrast to the disadvantages. Conjoint analysis requires a great deal of effort by researchers for design and analysis, as well as cognitive capacity from respondents. Two leading systematic reviews in the field (Ryan et al. 2001, Bridges et al. 2008) suggest that

research is needed into understanding the decision-making heuristics used when completing CA tasks. In the years since these reviews were published there has been research which suggests that during completion respondents violate the rule of additivity and compensatory decision-making (Gilbride & Allenby 2004) that CA is based on. Respondents instead appear to use a dual processing approach, where they use a combination of weighing factors systematically against each other (as in system 2) and making instinctive decisions (as with system 1) to simplify tasks (Stanovich & West 2000). Cunningham et al. (2010) suggest that Adaptive CBC allows for the simple non-compensatory decision-making processes to make tasks more acceptable to respondents. Respondents' decision-making heuristics have also been investigated and the assumptions of rationality and consistency questioned (Lancsar & Louviere 2006). Overall, presently the underlying economic modelling often does not reflect the capabilities of respondents (Payne et al. 1999) and there is still some way to go in order to validate the use of methods designed to help respondents to better manage CA choice tasks, such as ACBC (Cunningham et al. 2010).

Unlike marketing research, there is a distinct lack of external validity in health research. In other words, it is unclear whether what an individual states in a hypothetical clinical decision-making context accurately represents their real-life clinical decisions (Kievit et al. 2010). Only one study has found evidence of external validity in the context of prescribing decisions for alcoholism medication (Mark & Swait 2004). It has been suggested that this may be because it is more of a challenge to test external validity when a health service is publicly financed (Louviere & Lancsar 2009). External validity is perhaps more testable for private services and products, such as medications paid for over-the-counter, and this is

research which needs encouragement as it would add to the validation of stated preference techniques (Louviere & Lancsar 2009).

Moreover, the implications of CA results are limited to the attributes included in the design (Kievit et al. 2010) and because CA designs are constructed based on orthogonal designs this might mean that the attribute combinations are not entirely relevant to respondents. If a choice alternative, for example, does not include attributes salient to a respondent then they may not devote sufficient attention in order to fully assess their preferences (Cunningham et al. 2010).

2.4 Reasons for choosing CA to understand the decision to consult the GP for joint pain

Conjoint analysis has been chosen, rather than other approaches, to analyse determinants of GP consultation for joint pain for several reasons:

i. Unlike Likert scales, where respondents are able rate all attributes of a service or product as equal, CBC forces individuals to trade-off different attributes of the service, weighing up services that have more and less desirable characteristics (Phillips et al. 2002a). Focus groups suggest that CA encourages respondents to think more deeply about their preferences and better captures their choices in the real world (Phillips et al. 2002a).

ii. CA uses a 'decompositional' approach which utility to be estimated for each attribute level and for all possible combinations of levels, including combinations not directly evaluated by respondents (Phillips et al. 2002a).

iii. Epidemiological studies and qualitative studies tend to deal with observed patterns of use and personal experience of currently available services. Conjoint analysis allows utility estimation for any combination of attributes, including combinations that represent a health service that may not

currently be available. Patient preferences for these may inform future service development. This is particularly important for understanding the decision to consult the GP because of the need to develop optimal services and treatments (e.g. a new promising treatment) that may encourage those most in need of care (i.e. based on 'felt' need).

i. CA provides quantitative estimates of the relative importance of attributes; something which is usually unattainable in traditional observational and qualitative studies.

ii. The decision to consult is likely to be determined by a combination of need, enabling (including patient-centred aspects of care), and predisposing factors as well as outcomes of care. In theory, CA methods provide a theory-based framework for simultaneously investigating these. This thesis works from the premise that a more original contribution could be made by using CA techniques to complement the many such observational and qualitative studies that are already in print. Existing research findings may provide an excellent source of evidence for possible actionable attributes and levels.

iii. CA allows researchers to explore how the complex range of attributes that characterise the decision to consult the GP for joint pain interrelate in a continually changing process involving the internal weighing up of numerous considerations.

2.5 Statement of the thesis aims and objectives

The broad purpose of the thesis was to understand why some patients with joint pain/OA do not consult their GP despite apparent clinical need, and to

ascertain whether (changes in) certain aspects of service provision (would be expected to) facilitate consultation.

The primary aim was to quantify the relative importance of selected clinical need and general practice service factors ('attributes') in the decision to consult the GP in older adults with joint pain/OA.

Secondary aims were:

- To compare the direct ratings of the importance of attributes and the attribute importance scores from conjoint analysis by the same individuals.
- To identify different types of preferences within this heterogeneous population.

The specific objectives were:

- To use a combination of literature review, qualitative methods (focus groups, cognitive interviewing) and simulation studies to design a new conjoint analysis experiment based on a selection of attributes and levels judged salient to the decision to consult the GP for joint pain.
- To conduct, analyse, and interpret the above conjoint analysis experiment in a population-based sample of community-dwelling adults aged 50 years and over with joint pain/OA.
- To compare the rank order of attribute importance scores from conjoint analysis and the direct ratings of the importance of attributes.
- To use subgroup analysis to identify different types of consultation preferences amongst older adults with joint pain/OA and conduct between-group comparisons on selected demographic and other patient characteristics.

2.6 Structure of the thesis

This section details the structure of the thesis. This thesis structure is aligned to the recommended stages of CA development (as in Table. 2.3, on page 31).

Table 2.5: Synopsis of thesis chapters

Chapter One: Introduction and background to the determinants of general practice consultation for joint pain in older adults
• Outlines the importance of the decision to consult the GP for symptomatic OA, the complexities of the decision-making process itself and the methodological limitations and gaps in knowledge of existing methods.
Chapter Two: Conjoint analysis as an approach for investigating the determinants of GP consultation
• Outlines the potential value of conjoint analysis in determining issues of relative importance, in terms of the determinants of GP consultation.
Chapter Three: A systematic review of the previous applications of conjoint analysis techniques for eliciting preferences for patient-centred care in the GP consultation
• Identifies and describes previous published applications of CA techniques to elicit preferences for aspects of patient-centred care in the GP consultation. • Applies the recently published International Society for Pharmacoeconomics and Outcomes Research (ISPOR) checklist and critically reflects on the use of the ISPOR checklist for methodological quality.
Chapter Four: Methods I: developmental studies and preliminary testing of respondent efficiency
• Outlines development of CA questionnaire and aims to identify all relevant attribute levels as supported by evidence, to test that the construction of tasks were acceptable for respondents and to test that the data collection instrument was appropriate.
Chapter Five: Methods II: developmental studies: statistical efficiency
• Describes how the design was tested to ensure statistical efficiency. This relates to how precisely a given set of conjoint questions can estimate the parameters of interest prior to fielding the questionnaire.
Chapter Six: Methods III: Main study design, methods, and planned analysis
• Describes how the determinants of GP consultation for joint pain/OA were elicited: main study (quantitative) conjoint analysis experiment in a population-based sample of community dwelling adults aged 50 years and over with joint pain/OA.

Table 2.5: Synopsis of thesis chapters cont'd

Chapter Seven: Results I: survey response and descriptive characteristics of respondents to conjoint study
• Presents the flow of response to the study and compares respondents and non-respondents. • Presents the frequency of missing data within returned questionnaires and the descriptive characteristics of the conjoint analysis respondents.
Chapter Eight: Results II: direct rating of selected attributes, conjoint utilities and relative importance of attributes
• Identifies the relative importance of selected determinants (attributes) of the decision to consult the GP about joint pain. • Explores how attributes may be traded-off in the context of different clinically relevant scenarios. • Compares and contrasts the direct ratings of the importance of attributes and the attribute importance scores from conjoint analysis by the same individuals.
Chapter Nine: Results III: subgroup analysis
• Uses subgroup analysis to establish what factors influence patients' priorities in the decision to consult the GP for joint pain.
Chapter Ten: Discussion, conclusions and recommendations
• Discusses the overall findings from the PhD, main conclusions and implications for future research and practice.

3 Chapter Three: A systematic review of the applications of conjoint analysis techniques for eliciting preferences for patient-centred care in the GP consultation

3.1 Introduction

The popularity of conjoint analysis (CA) in healthcare lies in its ability to force individuals to make choices between healthcare products and services made up of various sets of attributes. Based on these choices researchers can ascertain the relative importance of different attributes of healthcare services as well as the trade-offs between these attributes presented at different levels (Ryan et al. 2001). To this end CA techniques have been used to analyse the factors associated with patient preferences across a variety of health services, including treatment preferences for osteoarthritis (Fraenkel et al. 2004a) and primary healthcare consultations (Cheraghi-Sohi et al. 2008). CA methods assume that healthcare services (like primary care consultations) can be described by a range of attributes and the value of a consultation to a respondent depends on the nature and level of these attributes. In this way researchers have been able to manipulate hypothetical attributes that are salient to patients' preferences. For example, enabling attributes such as location of treatment (Ryan et al. 2005) (process-based) and attributes related to patient-centred care such as 'doctor recognises your pain' (Haas 2005) have been analysed.

It appears from the literature review in chapter 1 that the decision to consult the GP for joint pain/OA is likely to be determined by a combination of need-related attributes (e.g. pain severity) and enabling factors (e.g. previous experience of GP consultation). Patient-centred care related attributes like GP attitude (i.e. patients feel that the GP does not take their OA seriously and that GPs ignore their symptoms) appear important in GP consultation for joint pain

(Alami et al. 2011), but may be under researched. At present, it is unclear whether CA has been used to capture clinical need related attributes in conjunction with service or patient-centred care related attributes in order to characterise such a complex decision (the primary aim of this thesis).

As part of the current thesis a systematic review of the literature was conducted, updating that of Ryan et al. (2001) and Bridges et al. (2008) but restricted to attributes of patient-centred care and the GP consultation. This was done in order to identify how previous studies have defined and described attributes that characterise patient-centred aspects of GP consultation. The attributes and levels and other common practices will be used to inform the attributes and levels used in the main study.

The Ryan et al. (2001) review of CA assessed its effectiveness for eliciting preferences from a non-specific population, and concluded that conjoint-based methods (including ranking, rating and choice-based) were recommended compared to other quantitative methods such as traditional direct ranking exercises (not conjoint based) whose results are of limited use (Ryan et al. 2001). However, they suggest that future research explore in more detail psychological issues related to completion of CA tasks, such as the cognitive strategies and heuristics adopted when completing decision-making tasks, i.e. using extensive piloting and developmental work using qualitative research to inform design and interpretation (Ryan et al. 2001).

Bridges et al. (2008) classified conjoint studies into three categories: clinical applications (n = 122), methodological contributions (n = 56) and health system applications (n=47) conducted between 1982 and 2007. The mean sample size among articles focusing on health system applications was 556 and the majority of

papers claimed to use orthogonal factorial designs, although over a quarter of papers did not report their design properties. Bridges (2008) argues that clearer guidelines for conducting and reporting conjoint analyses are needed (Bridges et al. 2008).

In the years since these reviews were published, more CA studies have appeared in the health literature as well as much work into cognitive strategies and heuristics. At the time of writing, one other systematic review of CA was ongoing and had published an abstract of their findings (Marshall et al. 2010). Marshall (2010) provided an update on current practice in the published literature between 2005 and 2008, identifying 79 conjoint analysis studies. Of these 71% used a discrete choice survey format, with the number of attributes ranging from 3 to 16. Most surveys included 6 attributes, and 73% presented 7–15 scenarios to each respondent. Sample size ranged from 13 to 1258 (mean sample size = 259) with most studies (38%) including between 100 and 300 respondents. Marshall (2010) concluded that there was large variation in methods, terminology, presentation and sample size, though the number of attributes, and number of scenarios presented to participants may assist researchers new to the field in developing their designs (Marshall et al. 2010).

While the stages involved in the design and analysis of CA studies have been described (see Table 2.3 on page 31), more guidelines for design and analysis are beginning to emerge (Bridges et al. 2011). In order to update the Ryan et al. (2001) and Bridges et al. (2008) reviews it was decided to use the checklist developed by Bridges et al. (2011). The recently published checklist provided by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) for *'undertaking, assessing, and improving the quality of*

conjoint-analysis applications in healthcare' (Bridges et al. 2009, p.2) (see Figure 3.1 on page 65) will be reflected upon critically.

3.2 Aims and objectives

The aims of this systematic review were:

1. To identify and describe previous published applications of CA techniques to elicit preferences for aspects of patient-centred care in the GP consultation.
2. To understand methodological strengths and limitations of previous studies.

Hence, the objectives were:

- To conduct a systematic search of the health literature for all published studies using conjoint analysis techniques for eliciting preferences for aspects of patient-centred care in the GP consultation and to extract and synthesise data relating to their design and analysis.
- To apply the recently developed quality checklist developed by the ISPOR special interest group to all studies included in the review.
- To highlight and discuss how published studies included in the review addressed or explored the attributes that characterise GP consultation.

3.3 Methods

3.3.1 Developing and designing the main search strategy

The unbiased and complete identification of studies is fundamental when compiling a systematic review (Higgins et al. 2008). Prior to conducting the main search, a series of steps were followed to develop and design the search strategy.

Establishing the remit of the review

The inclusion and exclusion criteria should define the study method, population, interventions and outcomes. The Population, Intervention, Comparison, Outcome (PICO) format was used to specify the remit of the review (see Table 3.1).

Table 3.1: PICO format used to specify the remit of the review

Participants, settings & conditions	Studies focused on primary care users (patients or consumers) and any health condition.
Interventions or exposures	Studies using conjoint analysis (CA) techniques, including choice-based conjoint (CBC), discrete choice experiments (DCE), adaptive conjoint analysis (ACA), and adaptive choice-based conjoint (ACBC)
Comparisons/ control	Not relevant
Outcomes of interest	Preferences for primary care consultation with emphasis on patient-centred care attributes (starting from continuity of care: your GP of choice: yes/no).

The scope of the review included patients' preferences for general practice and primary care consultation contexts only and excluded studies into secondary and tertiary care settings.

The review was intended to include all studies that had included attributes of patient-centred care, though these concepts were difficult to operationalise consistently. Patient-centred care attributes were defined according to the components defined by Mead and Bowers (2000) (see Table 3.2). Continuity of health professional (for example, getting to see your GP of choice) was identified as the starting aspect of patient-centred care (i.e. as an aspect of the therapeutic alliance in Table 3.2). Continuity of health professional is categorised as an aspect of patient-centred (access to) care by other researchers (Berry et al. 2003) and so attributes related to this were included in the review. Other attributes that were

considered as patient-centred care included ‘shared decision-making’. Papers with ‘technical care’ attributes or ‘process’ attributes only (e.g. perceived thoroughness of examination) were not included.

Table 3.2: Components of patient-centred care

Components of patient-centred (Mead & Bower 2000)
The **biopsychosocial** perspective: a perspective on illness that includes consideration of social and psychological (as well as biomedical) factors
The **patient as person**: understanding the personal meaning of the illness for each individual patient
Sharing power and responsibility: sensitivity to patients’ preferences for information and **shared decision-making**, and responding appropriately to these
The **therapeutic alliance**: developing common therapeutic goals and enhancing the personal bond between doctor and patient
The **‘doctor-as-person’**: awareness of the influence of the personal qualities and subjectivity of the doctor on the practise of medicine

Identifying key existing published review articles

Checks with the Cochrane Library were made to assess whether other systematic reviews of sufficient quality had already been conducted or were awaiting publication. This was supplemented by contact with key research groups within the field.

The background search used a range of synonyms for conjoint analysis techniques with no limitation on the type of the studies selected. The aims of the literature search were to locate and critically appraise as many published review articles as possible and identify the most popularly utilised conjoint techniques and all types of evidence relevant to the research question and to collect a wide range of information and terms/synonyms about conjoint analysis. Two relevant published systematic reviews (Ryan et al. 2001, Bridges et al. 2008). Bridges et al. (2008), one unpublished review (Marshall et al. 2010) and several articles recommended by experts, for example (Louviere & Lancsar 2009) were identified.

However, no published reviews were found that directly addressed the research aims of this thesis.

The citation lists of included studies were searched in order to identify any additional relevant studies.

Development of search terms

Developing the search strategy was an iterative process. A preliminary search of the literature, use of recent relevant published articles, and direct contact with leading experts in conjoint analysis research (not restricted to health fields) were used to develop potential search terms and keywords for the systematic search.

The final search was then based on interchanging preliminary terms and condensing the results using AND/OR functions. On the basis of the preliminary results, terms were removed that appeared irrelevant (e.g. cost benefit analysis) or generated vast numbers of papers (e.g. ACA/CBC abbreviations).

Initially a broad search strategy, including conjoint analysis terms only was conducted in order to identify an inclusive range of papers involving CA methods. However, this led to a high volume of papers (the majority of which were searched and examined) and so the strategy was developed by adding terms relevant to the GP consultation and patient-centred care. Searches were checked by assessing whether known key articles returned by the initial search strategy were identified. This led to the inclusion of some additional search terms, for example one identified study (Markham et al. 1999) identified ‘patient satisfaction’ as requiring inclusion in the final search strategy.

3.3.2 Main search strategy

Research question

The research question for the review was 'how have conjoint analysis techniques been used to investigate patient preferences for general practice consultation, what were their findings, and what were the strengths and limitations in this approach (including the validity of conjoint techniques in this context)?'

Sources

Electronic searches of the following databases hosted by EBSCO were undertaken in August 2009:

- Allied & Complementary Medicine (AMED) (1985 to August 2009).
- CINAHL (1982 to August 2009).
- PsycINFO: (1806 to August 2009).
- MEDLINE (1950 to August 2009).
- Academic Search Premier (1965 to August 2009). This database provides cited references for more than 1,000 titles.
- International Bibliography of the Social Sciences (1951 to August 2009).
- Ageline (1975 to August 2009).
- EMBASE (1974 to August 2009).

In addition to the health databases the following supplementary sources were searched:

- Websites of private companies providing CA software (Sawtooth Software and dobney.com).
- Selected experts and authors in the field of DCE and authors who had already published CA studies were also contacted.

- Reference and citation checking of key review articles and eligible articles identified by the health database search.

Search terms

A free text search was performed using all appropriate subheadings, thesaurus terms and MESH terms. All searches were limited to studies in humans. If appropriate to the database, truncation symbols (for example * used to represent 0 or more characters) were used. Final search terms and databases can be found in Appendix 1b (on page 319).

Selection of studies

Inclusion criteria were: original studies focused on preferences of patients or consumers with any health condition for general practice consultation; using conjoint analysis (CA) techniques, including choice-based conjoint (CBC), discrete choice experiments (DCE), adaptive conjoint analysis (ACA), and adaptive choice-based conjoint (ACBC); including one or more patient-centred care attribute; full article, English language. Studies that were non-human, or not primary studies (e.g. editorials, letters, methodological reviews) were excluded.

In phase one of exclusion, the titles and abstracts of all identified articles were screened by one reviewer (DC) and any studies not meeting the criteria outlined in Table 3.1 were excluded. To verify consistency in applying the eligibility criteria the abstracts of 10 randomly selected papers were reviewed for content by a second independent reviewer (BA-O). Any queries or uncertainties were passed to a third reviewer (GP) for arbitration before a final decision agreement was made by the first reviewer (DC).

Next, the full text articles of all potentially eligible titles/abstracts were checked by the first reviewer (DC) for eligibility. The remaining eligible papers were then grouped according to the objectives of the review (see section 3.2 on page 57).

Data extraction

The following data was extracted from all eligible papers by a single reviewer (DC).

- Place of study
- Total number of participants (response rate)
- Sampling
- Setting
- Participant characteristics
- Conjoint analysis technique
- Experimental design
- Main findings
- Statistical analysis procedures

Information on tests conducted for reliability, validity, rationality and their results were separately extracted.

Quality assessment

The quality of the selected studies was assessed by two independent reviewers (DC, GP), using the recently published ISPOR checklist which comprised of 30 items (see Figure 3.1 on page 65). Although it had not been formally evaluated it promised nevertheless a clear starting point in the assessment of conjoint analysis studies.

For each of the 30 items in the ISPOR checklist the items were scored as yes (Y), no (N) or unclear (?). No studies were excluded from the analysis on the basis of quality assessment.

The degree of agreement between the readers was calculated using a Kappa coefficient - a measure of how much better the researcher agreement was than if it were due to chance. This offers a more robust assessment than a simple percentage agreement calculation (Altman 1991). The final agreed scores were recorded using a quality assessment table (see Appendix 2, on page 329).

Figure 3.1: ISPOR checklist items (Bridges et al. 2011)

ISPOR checklist item		Items	
1.	Was a well-defined research question stated and is CA appropriate?	**1.1**	Was a well-defined research question and/or testable hypothesis articulated?
		1.2	Was the study perspective described and the study placed in any particular decision- making or policy context?
		1.3	What is the justification for using conjoint analysis to answer the research question?
2.	Were the attributes and levels supported by evidence?	**2.1**	Were all important and relevant attributes identified (i.e. supported by literature reviews/focus groups/other methods)?
		2.2	Was the choice of included attributes justified and consistent with theory?
		2.3	Were the range, description and number of levels for each included attribute justified?
3.	Was the construction of conjoint tasks appropriate?	**3.1**	Was the number of attributes in each conjoint task justified (that is, full-profile or partial-profile)?
		3.2	Was the number of profiles or alternatives in each conjoint task justified?
		3.3	Was the number of conjoint tasks included in the data- collection instrument appropriate?
4.	Was the choice of experimental design justified and evaluated?	**4.1**	Was the choice of experimental design justified? Were alternative experimental designs considered?
		4.2	Were the properties of the experimental design evaluated?
		4.3	Was (should) an opt-out or a status-quo alternative (be) included? Conducted appropriately? Were issues of clustering and sub- groups handled appropriately?
5.	Were preferences elicited credibly?	**5.1**	Were the conjoint tasks sufficiently motivated and explained?
		5.2	Was an appropriate elicitation format (rating, ranking, or choice) used? Did (should) the format allow for indifference?
		5.3	In addition to preference elicitation, did the conjoint tasks include other questions (e.g. strength of preference, confidence in response, and other methods?)
6.	Was the data-collection instrument designed appropriately?	**6.1**	Was appropriate information about respondents collected (sociodemographic/attitudinal/health history/status/treatment)?
		6.2	Were the attributes adequately described and was necessary contextual information provided?
		6.3	Was the level of burden of data-collection instrument appropriate? Were respondents informed/encouraged/motivated?
7.	Was the data collection plan appropriate?	**7.1**	Was the sampling strategy justified (for example, sample size, stratification, and recruitment)?
		7.2	Was the mode of administration justified and appropriate (for example, face-to-face, pen-and- paper, web-based)?
		7.3	Were human-subjects considerations addressed (e.g. recruitment, information and/or consent, compensation)?
8.	Were statistical analyses and model estimation appropriate?	**8.1**	Were respondent characteristics examined and tested?
		8.2	Was the quality of the responses examined (for example, rationality, validity, reliability)?
		8.3	Was multivariate analysis conducted appropriately? Were issues of clustering/sub- groups handled appropriately?
9.	Were results and conclusions valid?	**9.1**	Did results reflect testable hypotheses and account for statistical uncertainty?
		9.2	Were conclusions supported by the evidence and compared to existing findings in the literature?
		9.3	Were study limitations and generalisability adequately discussed?
10	Was the study presented well and completely?	**10.1**	Was the study importance and research context adequately motivated?
		10.2	Were the study methods explained and the data- collection instrument adequately described and/or illustrated?
		10.3	Were the implications of the study stated and understandable to a wide audience?

Synthesis of Material

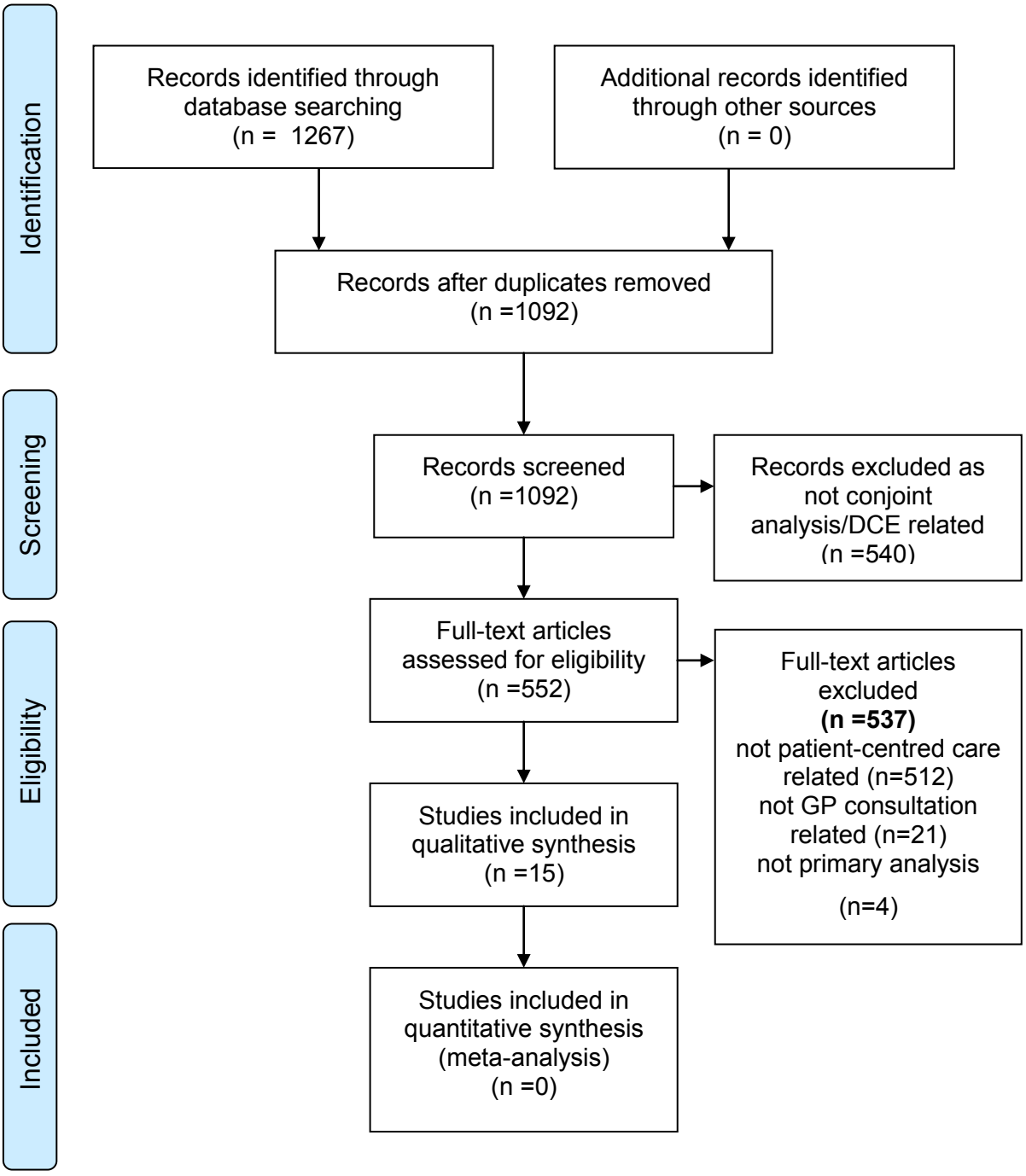
Studies were evaluated according to the content of the ISPOR checklist that consists of 10 key items (see Figure 3.1). Key references such as Ryan et al. (2001) and Ryan and Gerard (2003) were used in addition to other existing published systematic reviews of conjoint analysis (Bridges et al. 2011, Marshall et al. 2010).

3.4 Results

3.4.1 Studies included in the review

A total of 1267 citations were identified (MEDLINE 356, CINAHL 147, PsychINFO 284, EMBASE 60, Amed 17, Ageline 37, Academic Search Premier 261, International Bibliography of the Social Sciences 105) and 175 duplicates were excluded. The remaining 1,092 abstracts were read, and a further 1,077 studies were excluded based on exclusion criteria. The main reasons for exclusion were that studies did not use conjoint analysis methods (n=540) and did not focus on patient-centred care (n=512) in the context of GP/primary care consultation (n=21). If studies were not primary empirical studies (n=4), such as methodological reviews, these were excluded.

Figure 3.2: Flow diagram of study selection (Moher et al. 2009)



3.4.2 Description of studies

This systematic review found 15 primary studies that used conjoint analysis techniques for eliciting preferences for aspects of patient-centred care in the GP consultation. All 15 studies were in a general practice or primary care setting. Eleven were based in the UK, two in Australia, one in Sweden and one in the USA. Final sample size ranged from 51 to 3,893 participants with a median of 666. Response rates ranged from 18% to 94% with a median of 60.5% (based on fourteen studies, as one study did not report a response rate). All study samples consisted mostly of females and ages ranged from 16 to 86.

Table 3.3: Data extraction table

Studies		Place	n (response rate, %)	Sampling	Setting	Participant characteristics	CA type	Experimental design	Main Findings	Statistical analysis
1	Caldow et al. (2007)	UK	1343 (49%)	Patients registered with general practices. Randomised postal questionnaire survey followed by voluntary telephone interviews	Primary care	Age 16-75+ 59% F	DCE	5 attributes.15 levels. 8 choice sets. Two alternatives and an 'opt-out'	Though most people would prefer to see a doctor if nurses took on more roles traditionally associated with doctors, patients would accept them	Nested logit regression
2	Cheraghi-Sohi et al. (2008)	UK	1193 (53%)	Patients registered with general practices. Randomly selected and stratified by sex and 3 age bands. Postal: 1 of 24 different questionnaire versions	Primary care	Generic que: Mean age 52.5 years SD:17.9 55% F Patient-centred care que: Mean (SD) 53 years (18.3) 52% F	DCE	2 4-level attributes and four 2-level attributes. 16 choice sets with 2 alternatives, blocked into 2 sets of 8 choice sets	Although patient-centred care is important to patients, they place more importance on continuity of care and technical care	Probit and Random effects Probit
3	Fiebig et al. (2009)	Australia	167	Stratified random sample aged 18–69 who previously had Pap test. Postal	General practice	18–69 years 100% F	DCE	6 ''context'' and 3 ''alternative specific'' attributes. 32 choice sets with 3 alternatives	In the decisions about pap tests women prioritised cost, chance of a false positive and if recommended screening interval was 1 year	Mixed logit Multinomial logit
4	Gerard et al. (2008)	UK	1052 (94%)	Patients registered with general practices: patients attending GP/nurse consultation under 16 were handed questionnaires	General practice	Median age 52 years 64% F	DCE	4 attributes, 11 levels. 8 choice sets (in two different questionnaires) 2 alternatives	Seeing a doctor of choice, time of day of appointment, number of days wait for an appointment and length of appointment were prioritised	Conditional logit model

Table 3.3 cont: Data extraction table

Studies		Place	n (%)	Sampling	Setting	Participant characteristics	CA type	Experimental design	Main Findings	Statistical analysis
5	Haas. (2005)	Austra-lia	128 (88 %)	Randomised from patients who visited GP in the last 6 months for treatment of minor conditions Market researchers recruited	General practice	18 - 80 years Mean age 47.8 years 53% F	DCE	7 attributes (18 levels). 24 scenarios viewed. 3 alternatives	Trust, legitimation, doctor recognises your pain/distress, doctor treats you with dignity, doctor reassures you, doctor provides information, doctor accepts your decisions about your health were attributes valued most highly	Multinomial logit regression
6	Hjelmgren et al. (2007)	Sweden	924 (58%)	Postal questionnaires to 18 and 85 years - stratified to the three largest cities and the rest of the country	Primary care	Mean age 48.9 years 59% F	DCE	5 attributes, 13 levels. 16 choice combinations randomly split into 4 choice sets. Two alternatives	Patient influence over the care received was valued most highly and the primary care work model was the least	Random effects logit regression.
7	Longo et al. (2006)	UK	584 (78%)	Postal to RCT patients 6 months later. Patients had (atrial fibrillation, menorrhagia, menopausal symptoms, prostatism)	General practice	Mean (SD) age 59 years (10.7) 54% F	DCE	5 attributes, 12 levels. 24 comparisons divided between questionnaires. 14 pairwise choices presented	Larger utilities in 'doctor listens' attribute, followed by easily understood information, a shared treatment decision, more information and longer consultation	Multi-level logistic regression
8	Markham et al. (1999)	USA	223 (76%)	Random convenience sample approached in GP waiting room. Student assistant interviewed patients	Primary care	Median age 37 years 75% F	CA rating	5 attributes and 11 levels. Fractional-factorial design (8 scenarios presented)	Perceived skill of the physician was most important and time in the waiting room least important	Mixed and Multinomial logit model

Table 3.3 cont: Data extraction table

Studies		Place	n (%)	Sampling	Setting	Participant characteristics	CA type	Experimental design	Main Findings	Statistical analysis
9	Morgan (2000)	UK	271 (65%)	Random sample of respondents who consented to further contact to a previous Sheffield-based survey into out-of-hours care. postal surveys	Primary care	16 – 95 years Median age 50 years 54% F	DCE	6 attributes, 16 levels 16 pairwise choice sets	Doctor's manner (takes time to listen), type of consultation, and waiting time for consultation were most important	Probit regression analysis
10	Rubin et al. (2006)	UK	1153 (55%)	Random selection from general practices. questionnaires handed out to patients over 18 years attending appointments	General practice	18- 90 years Mean age 46.15 years 67% F	DCE	3 attributes, 8 levels. 7 choice sets. Two alternatives	Choice of doctor, Time to appointment and choice of time were the most valued attributes	Probit regression models; fixed and random effect
11	Ryan et al. (1998)	UK	51 (51%)	Random selection of people aged 16 years + from an evaluation database of an RCT were sent postal questionnaires	Primary care	16 years + unreported % F	CA rating	4 attributes, 9 levels. Fractional-factorial design (8 scenarios)	The patient health card is least significant. The no of days between making a non-urgent appointment and seeing a doctor were most important	Ordinary least squares regression
12	Scott and Vick (1999)	UK	734 (18%)	Random sample of patients registered with general practices. Postal	General practice	Age range 16 – 70 years 57% F	DCE	5 attributes, 12 levels. 2 alternatives. (8 choice sets)	Being able to talk to the doctor was most important, followed by patients understanding of doctor's explanation	Multilevel and random effects probit regression

Table 3.3 cont: Data extraction table

Studies		Place	n (%)	Sampling	Setting	Participant characteristics	CA type	Experimental design	Main Findings	Statistical analysis
13	Scott et al. (2003)	UK	3893 (68%)	Postal questionnaire sent to parents of children in Aberdeen and Glasgow who had received a home visit or attended a primary care emergency centre or were registered with a GP	General practice	16-75 years Mean age 34years 87% F	DCE	4 attributes, 11 levels. 8 pairwise choice sets	Those who had never used out-of-hours care before had stronger preferences for waiting time and the doctor listening, suggesting higher expectations of non-users	Random effects probit regression
14	Turner et al. (2007)	UK	646 (47%) postal + 20 interviews (n=666)	Age stratified sample from GP practices. Postal and face-to-face interviews	General practice	Mean age 58 years 61% F	DCE	4 attributes, 10 levels. 7 pairwise choice sets	Patients would wait longer to see a familiar medical practitioner who was well informed about their case when they had a problem causing uncertainty or needed a routine check up	Random effects probit regression
15	Vick and Scott (1998)	UK	101 (63%)	Prepaid envelopes were handed out to attendees of a GP practice (parents and guardians)	General practice	Mean age 36.4 years 73% F	DCE	6 attributes, 14 levels. 13 pairwise choice sets	Being able to talk to your doctor was most important while who chooses your treatment was least important	Random effects Probit regression

3.4.3 Narrative synthesis: study findings

The majority of attributes used in the included CA studies were enabling factors as defined in the Andersen-Newman model (Andersen 1995). None of the studies included clinical need factors. One study used a health outcome related attribute; likelihood of having illness cured (Caldow et al. 2007). In the table in Appendix 1c attributes from the reviewed studies were further categorised as patient-centred care, interpersonal, technical or process attributes in primary care. Attributes were ranked in order of importance from 1 – 7 (i.e. the highest amount of attributes in one study) in order to synthesise their importance across studies (see Appendix 1c on page 323).

The attribute most frequently ranked as the most important across studies was whether the 'doctor listens' (an aspect of interpersonal care), appearing in five studies (Vick & Scott 1998, Scott & Vick 1999, Morgan et al. 2000, Scott et al. 2003, Longo et al. 2006). Typically in these studies, 'being able to talk to the doctor' was more important than other process related attributes such as waiting times, and the type of explanation received. This was evident in a small survey (n=101 patients), (Vick & Scott 1998) a larger postal survey (n=734) (Scott & Vick 1999) and in studies into out-of-hours care (Morgan et al. 2000, Scott et al. 2003).

However, although only two studies included technical care attributes, in these studies technical care was found to be the most important attribute, compared with interpersonal and patient-centred aspects of care (Markham et al. 1999, Cheraghi-Sohi et al. 2008). This may support the argument that unless these attributes are specified patients may tend to assume that technical care is of a good quality (Chapple et al. 2002). Cheraghi-Sohi et al. (2008) specifically selected attributes of technical and patient-centred care in order to define the

relative importance of these in primary care consultations. They concluded that patients were willing to pay the most for a thorough physical examination, with attributes 'seeing a physician who knew them well', 'seeing a physician with a friendly manner' and 'having flexible appointment times' being less important.

The second most important attribute across studies was 'seeing a doctor of choice' (or a doctor who knew them), appearing as most important in three studies (Rubin et al. 2006, Cheraghi-Sohi et al. 2008, Gerard et al. 2008). In one study speed of access was far outweighed by accessing their own GP (Rubin et al. 2006). However, in one study the 'patients influence over the care received' was more important than 'choice of GP' (Hjelmgren & Anell 2007). In another two studies the type of professional (GP or nurse) was more important than the continuity of professional (Caldow et al. 2007, Turner et al. 2007). Thus, across reviewed studies continuity of health professional was not as important as whether or not they listen.

The third most important attributes across studies were process related, appearing as most important in three studies (Ryan et al. 1998, Caldow et al. 2007, Turner et al. 2007). They appeared as second most important in seven studies and third most important in eight studies (see Appendix 1c on page 323). Waiting time until appointment (e.g. no waiting time, 2 days, 4 days, 8 days) appeared to be most commonly included in studies. Other process attributes included 'choice of time' (your choice of time or at a specified time) (Rubin et al. 2006) and information about your health problem from the doctor (Vick & Scott 1998).

Six studies included attributes related to shared decision-making, a specific aspect of patient-centred care (Vick & Scott 1998, Markham et al. 1999, Scott &

Vick 1999, Haas 2005, Longo et al. 2006, Cheraghi-Sohi et al. 2008). Shared decision-making attributes appeared to be less important than whether the doctor listens or the choice of health professional. For example, Longo et al. (2006) found that although patients with chronic conditions valued shared decision-making, it was less important than the doctors' ability to listen. Markham et al. (1999) also found that shared decision-making along with process based aspects of care like waiting time and billing problems was less important in determining patient choice of doctor, than technical care.

In summary, the most important attributes across studies were whether the doctor listens (interpersonal attributes) followed by the choice of health professional (an aspect of patient-centred care). Process attributes were also important but shared decision-making attributes appeared to be least important across studies. When technical care attributes were included in studies (which happened rarely) they appeared to be the most important. However, it is important to note that although the importance of attributes across studies can be synthesized to give an indication of the importance of attributes, it is clear that the relative importance scores are very task and context-dependent (i.e. dependent upon which combination of attributes was included in the CA tasks).

3.4.4 Narrative synthesis: quality assessment

Inter-rater agreement on ISPOR checklist

Table 3.4 (on page overleaf) shows the classification by two researchers of the ISPOR checklist items of 15 studies as 'Yes' 'No' or 'Unclear'. See Appendix 2 for results of independent quality assessment by two reviewers. Both reviewers rated the majority of sub-items across the 15 studies as positive, although reviewer B was generally more negative while reviewer A was more likely to rate a

sub-item as 'Unclear'. Further detailed results on the independent quality assessment by the two reviewers and the inter-rater agreement are provided in Appendix 2 (page 329).

Agreement was observed on 77% of the sub-items. However, this does not take into account the possibility that some of this agreement was due to chance.

Table 3.4: Classification of the ISPOR checklist

	Reviewer B			
Reviewer A	YES (%)	NO	UNCLEAR	Total
YES	271	87	38	396
NO	4	33	2	39
UNCLEAR	3	3	9	15
Total	278	123	49	450
Observed agreement: 304/395 = 77% Unweighted Kappa (95%CI) = 0.32 (0.20, 0.44)				

To account for this chance agreement, an unweighted kappa, (κ) was calculated. Kappa was only performed on Y/N responses. Kappa has a maximum of 1 when agreement is perfect, zero indicates no agreement better than chance and negative values show worse than chance agreement. While the interpretation of kappa values is context-dependent there are general guidelines (Altman 1991).

Table 3.5: Guidelines of inter-rater agreement by Altman (1991)

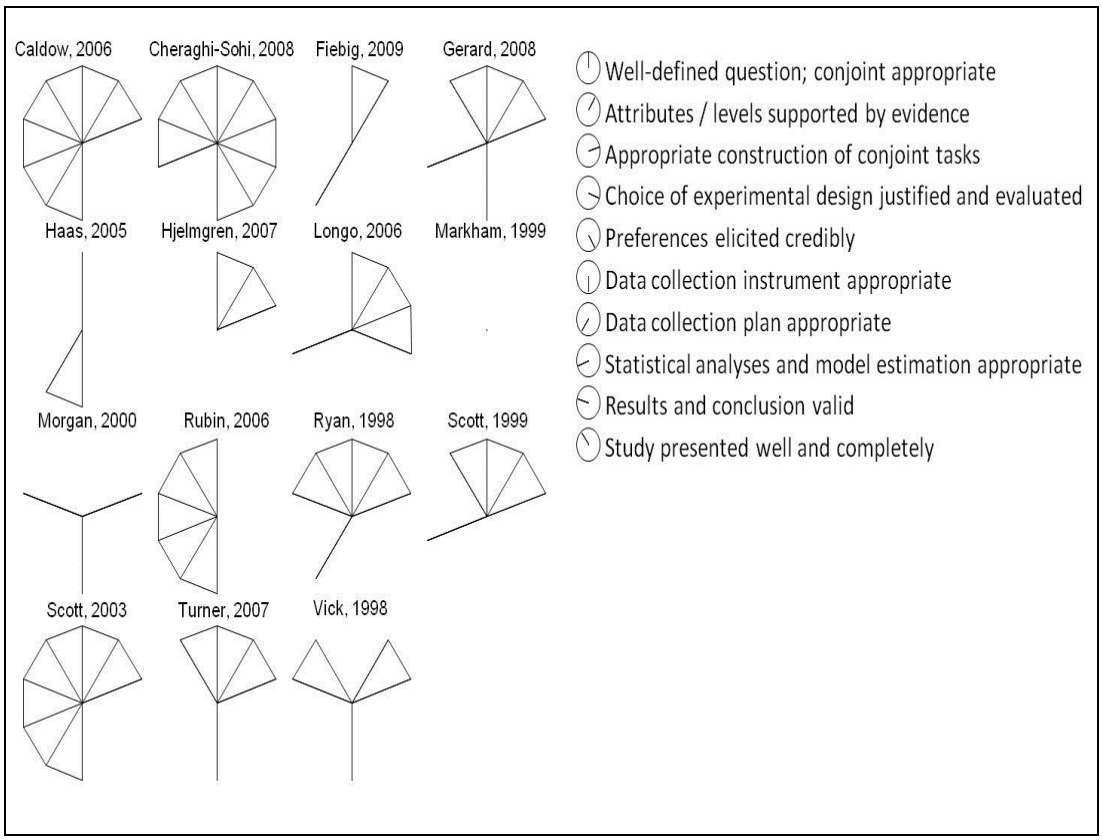
Value of Kappa	Strength of agreement
< 0.20	Poor
0.21 – 0.40	Fair
0.41 – 0.60	Moderate
0.61 – 0.80	Good
0.81 – 1.00	Very good

It can therefore be said that there was fair agreement between the two raters. Item 1 (was a well-defined research question stated and is conjoint an appropriate method?) had the highest observed agreement (82%) - and item 4

(was the choice of experimental design justified and evaluated?) had the lowest agreement (49%). The second lowest observed agreement was with item 3 (62%) (construction of conjoint tasks). In both items 3 and 4 researcher B was a little more cautious (rating 'No') than researcher A (rating 'Yes').

The following section takes a narrative approach to the quality appraisal of studies, based on a pictorial representation where each star represents one study (see Figure 3.3). Each spoke denotes satisfactory quality on each of the 10 ISPOR checklist domains. It was decided that satisfactory quality for a specific domain would be determined by a positive response for all items in that domain (i.e. only three Y's = yes, any other combination of responses = no/unclear. The results are reported under subheadings that reflect the five steps of design and analysis of a CA study (from Ryan & Gerard 2003) (see Table 2.3 on page 31).

Figure 3.3: Pictorial representation of methodological quality of reviewed studies based on agreed consensus of two reviewers



Stage 1. Selection of salient attributes, levels, and scenarios

A clear research question or hypothesis typically involved a clear statement of the objectives of the study. For example, *'this study was designed to investigate patients and prospective patients' attitudes to and preferences for changes in the delivery of primary healthcare, with particular reference to an extended role for practice nurses'* (Caldow et al. 2007, p 31). Two studies failed to express an aim or hypothesis (Markham et al. 1999, Morgan et al. 2000) and another study was unclear (Vick & Scott 1998). The majority of studies placed their study in a particular decision-making or policy context.

The reviewed studies reported a range of approaches to identifying salient attributes. These included undertaking preliminary qualitative work (Cheraghi-Sohi et al. 2008, Caldow et al. 2007, Morgan et al. 2000, Scott et al. 2003, Rubin et al. 2006, Gerard et al. 2008, Turner et al. 2007, Ryan et al. 1998), including focus groups (Caldow et al. 2007, Morgan et al. 2000) and discussions (Rubin et al. 2006), literature reviews (Markham et al. 1999, Morgan et al. 2000, Rubin et al. 2006), asking the opinion of clinicians (Ryan et al. 1998), drawing on the experience of investigators themselves (Markham et al. 1999), or a combination of these. Few studies piloted questionnaires (Scott et al. 2003, Longo et al. 2006, Turner et al. 2007) to allow researchers to refine the attribute levels. One study used 'think-aloud' cognitive interviews to verify the use of words used (Cheraghi-Sohi et al. 2008).

There were a variety of attributes used in the studies, though a few attributes were common across study contexts (e.g. 'doctor listens') (see Appendix 1c). 'Technical skill of doctor' appeared in one study as 'good medical skills' (Markham et al. 1999) and in another as 'thoroughness of care' (Cheraghi-Sohi et

al. 2008). Patient-centred care attributes such as continuity of care were either dichotomised (e.g. your choice of doctor or any available, Rubin et al. 2006), mixed with type of professional (e.g. Gerard et al. 2008) or were dichotomised using more emotive language (e.g. physician's knowledge of patient as in Cheraghi-Sohi et al. 2008). Contrary to suggestions in CA literature (Johnson 2008), CA methods appear to extend successfully to abstract concepts such as qualitative, patient-centred aspects of consultation in healthcare (as successfully as familiar products used in marketing research). Three studies measured willingness to pay using a cost attribute (Cheraghi-Sohi et al. 2008, Hjelmgren & Anell 2007, Fiebig et al. 2009), although it is unclear how relevant this is to UK-based studies since healthcare is free in the UK.

The description of the process of specifying attribute levels was not explicit in most studies (a finding consistent with Marshall et al. 2010). Turner et al. (2007) reached a consensus on five attributes to be included in pilot work, and after piloting only four were included (Turner et al. 2007). One study included attributes where the majority had two levels per attribute (Vick & Scott 1998). However, it is accepted that some dichotomous attributes will need to be simplistic in order to be correct e.g. usual doctor (yes/no).

Stage 2: Experimental design and construction of choice sets

All studies were considered to have constructed tasks appropriately i.e. they used a fractional-factorial design, using computer software to reduce the number of choice sets in order to make the task manageable for respondents. None of the studies considered using a partial-profile design. One study made a sensible set of decisions to exclude certain combinations (Rubin et al. 2006), though it was unclear whether orthogonality and balance was maintained. A d-

efficiency/optimality statistic was calculated in three studies to maximise the efficiency of the design (Longo et al. 2006, Cheraghi-Sohi et al. 2008, Gerard et al. 2008). The D-efficiency assesses the efficiency of the design based on how well it fulfils the properties of orthogonality (attribute levels are uncorrelated across consultation alternatives), level balance (attribute levels occurs an equal number of times) and minimal level overlap (attribute levels appear once in a choice alternative). These properties were introduced on page 33 in Chapter 2. The aim is for the efficiency of the design to be as large as possible (Huber & Zwerina 1996).

Thirteen studies used a discrete choice experiment format (87%), with number of attributes ranging from 3 to 7 and the number of levels ranging from 2 to 5. The majority of surveys included 5 or 6 attributes, and the number of choice sets presented to respondents ranged from 7 (Caldow et al. 2007) to 32 (Fiebig et al. 2009). Eight scenarios were presented in both of the CA rating studies. Another study presented 24 choice sets, double the 12 sets recommended by Ryan et al. (2001) (Haas 2005). In the thirteen (87%) that were DCEs ten (77%) were pairwise comparisons (consisting of two alternatives/profiles). One study included an opt-out (i.e. reject both) option with two alternatives (Caldow et al. 2007) - claiming to add to the realism of choice tasks - and two included three alternatives (Haas 2005, Fiebig et al. 2009).

Only one study justified their choice of method (Ryan et al. 1998) arguing that despite the ability for a DCE format to imitate real-world choice behaviour, rating data can be analysed as easily and requires further exploration. One study used an alternative-specific design, where some attribute levels are appropriate for some choice alternatives but not others (Fiebig et al. 2009).

Stage 3. Measurement of preferences

In all studies it was not clear exactly what literature accompanied the experiments, or whether conjoint tasks were sufficiently motivated and explained. In order to understand the motivation of respondents response rates were observed. Fourteen studies had acceptable response rates and one study had a very low response rate (18%) (Scott & Vick 1999). Seven studies included a difficulty rating and most of these studies reported that the majority of respondents did not struggle with the difficulty of the task. For example, 80% found the questionnaire ‘OK’ ‘quite easy’ or ‘very easy’, 19 % ‘found it quite difficult’ and 1% found it ‘very difficult’ (Scott & Vick 1999). However, in this study this is not indicative of ease of task since the reasons for non-response were rarely reported in any great detail.

Studies used two different CA techniques: discrete choice experiments and rating CA. Thirteen studies used a discrete choice experiment format (87%), with 77% of those offering a constrained choice between A and B (which gives no option of saying 'either', so therefore leaves no place for indifference). Two (13%) used CA rating techniques (Ryan et al. 1998, Markham et al. 1999). In rating CA individuals were presented with each of the scenarios and asked to state their level of preference for each scenario on a scale of one to five, where one indicates ‘dislike very much’ and five indicates ‘like very much’.

In six studies it was not clear exactly whether the attributes were adequately described and necessary contextual information provided as a template of the document given to respondents was not included. The majority of the studies collected the appropriate information about respondents (sociodemographics such as age and gender).

Respondent burden was mostly considered in terms of the amount of choice sets presented. The lowest number of choices presented to respondents was reduced to 7 (Rubin et al. 2006) and the highest was 32 (Fiebig et al. 2009). In line with guidelines, any more than 16 choice sets was considered too much of a burden to respondents (Haas 2005).

Final sample size ranged from 51 to 3,893 participants with a median of 666. Response rates ranged from 18 to 94% with a median of 60.5%, suggesting that tasks did not cause too much confusion or cognitive burden for respondents. Four of the studies had seemingly inadequate (small) sample sizes (Vick & Scott 1998, Ryan et al. 1998, Haas 2005, Fiebig et al. 2009). The Fiebig et al. (2009) study had a complex design that was heavily weighted towards statistical analysis but lacked basic results such as response rate and descriptive characteristics of respondents, though there was supplementary data and questionnaire available on request from authors.

Missing data was reported in four studies (Cheraghi-Sohi et al. 2008, Scott et al. 2003, Turner et al. 2007) and non-response was rarely reported (Ryan et al. 1998, Longo et al. 2006, Cheraghi-Sohi et al. 2008, Gerard et al. 2008). Studies suggested that patients who were older, female (Cheraghi-Sohi et al. 2008) did not care for someone at home, used a car to get to the surgery, were unemployed and were attending the surgery for an ongoing medical problem were more likely to respond (Gerard et al. 2008).

The majority of studies randomly selected respondents from patients registered with general practices. Others were stratified to the three largest cities or the rest of the country (Hjelmgren & Anell 2007). Ethical approval was stated in all 15 studies. All 15 studies were paper-based and there was little consideration

of other methods (i.e. computer-based). The majority of studies used self-complete postal questionnaires, with one recruiting six months from the start of an ongoing randomised control trial (Longo et al. 2006) and another doing telephone interviews with those who consented (Caldow et al. 2007). Recruitment involved written reminders (Longo et al. 2006), market researchers (Haas 2005) and recruiting patients waiting for appointments (Vick & Scott 1998, Markham et al. 1999, Rubin et al. 2006, Gerard et al. 2008).

Stage 4. Estimation procedure

The most commonly used analytic technique was a random effects probit model (Vick & Scott 1998, Scott et al. 2003, Turner et al. 2007). Other methods used included multinomial logit (Haas 2005, Fiebig et al. 2009) nested logit regression (Caldow et al. 2007), ordinary squares regression (Ryan et al. 1998), probit regression (Morgan et al. 2000, Rubin et al. 2006), multilevel logistic regression (Scott & Vick 1999) and conditional logit (Gerard et al. 2008).

Stage 5. Tests of the validity of responses

Most studies used tests of respondent consistency to assess whether respondents always chose the same alternatives when presented with identical choices (Cheraghi-Sohi et al. 2008). Rationality was examined using dominant options (options that were superior on all levels of attributes and therefore expected to be chosen) in some studies (Vick & Scott 1998, Longo et al. 2006, Caldow et al. 2007). Few of these studies excluded irrational responses from analysis (Longo et al. 2006). In one study it was considered that irrational responses did not influence results dramatically (Rubin et al. 2006). None of the studies tested the validity of models estimated from experimental data against real

market choices. Another aspect of validity was tested in fourteen studies by comparing their findings with what might be predicted by existing theory. In one study test–retest reliability was measured after the focus groups by asking the 46 participants invited, to complete the questionnaire twice, 3 weeks apart (Caldow et al. 2007). One study compared DCE estimates with a rating scale (5 point: 0= not very important and 5= very important) (Hjelmgren & Anell 2007). They concluded that DCEs allow researchers to be more precise about the trade-offs, as opposed to direct ratings which allow all attributes to be considered equal.

Study limitations and generalisability were adequately discussed. In some studies conclusions were not supported by evidence and compared to existing findings in the literature (Markham et al. 1999) while other studies did not discuss selection bias adequately (Scott & Vick 1999). All 15 studies state which attributes/levels included in the tasks were (or were not) significant and report uncertainty (relevant to the models they used). A range of confidence intervals, standard errors, coefficients and p-values were presented.

Some studies failed to explain/illustrate the methods and the data collection instrument adequately (Markham et al. 1999, Ryan et al. 1998, Scott & Vick 1999, Longo et al. 2006, Hjelmgren & Anell 2007, Fiebig et al. 2009). For example, many studies failed to give an accurate reproduction of the actual task instructions presented to respondents. This was a finding consistent with Marshall et al. (2010). Overall, the implications of the studies were stated and understandable to a wide audience in all studies, with one study lacking clarity (Haas 2005).

3.5 Discussion

3.5.1 Principal findings

The most important attributes across studies was whether the doctor listens followed by the choice of health professional, process attributes and shared decision-making. When technical care attributes were included in studies they appeared to be the most important. Thirteen studies used a discrete choice format (87%), with the majority of these offering a constrained pairwise choice between A and B. Attributes ranged from 3 to 7 and the number of levels ranged from 2 to 5. The majority of surveys included 5 or 6 attributes, and the number of choice sets presented to each respondent ranged from 7 (Caldow et al. 2007) to 32.

Overall, choice of attributes had justification from literature but whether these were systematic, comprehensive or impartial is unclear. Seven studies used qualitative research in conjunction with a literature review. All studies used a fractional-factorial design, using computer software to reduce the number of choices presented to respondents. The response rates overall were reasonable, despite not finding any evidence of direct motivation and encouragement to respondents.

Respondent consistency and dominance testing was used to test for consistency and rationality in studies though some studies did not exclude these responses from analysis. The majority of studies looked at another aspect of validity which compared their findings with what might be predicted by existing theory. One study tested for test–retest reliability. One study compared DCE estimates with a rating scale. None of the studies tested for external validity.

3.5.2 Interpreting the principal findings in the context of previous studies

a) What types of CA have been used?

The popularity of the discrete choice format (87%), including constrained pairwise choice tasks between (A and B) is consistent with an unpublished systematic review of CA (Marshall et al. 2010) where 71% of studies used a choice format. Two studies used CA rating techniques. Choice-based tasks are the least demanding for respondents cognitively, though they are criticised for not providing as much preference information and commonly rating and ranking methods are used alongside them (Flynn et al. 2007).

Attributes ranged from 3 to 7. In the most recent systematic review attributes ranged from 3 to 16 (Marshall et al. 2010). In this review the majority of surveys included 5 or 6 attributes, something comparable to the Marshall (2010) review. The number of choice sets presented to each respondent ranged from 7 (Caldow et al. 2007) to 32 (Fiebig et al. 2009). Thus the range of number of choice sets included in studies went above the recommended number of 12 (Ryan et al. 2001).

b) The methodological quality of the included studies

The methodological quality of the published studies was assessed using the ISPOR checklist, and varied across studies. Overall, choice of attributes had justification from literature (as the minimum requirement) but in almost all studies this literature review was not necessarily a systematic, comprehensive or impartial review of the literature. This was a finding consistent with the latest review in this field (Marshall et al. 2010). The result is that the choice of attributes might reflect the particular interests and biases of the researchers.

Seven studies used qualitative research in conjunction with a literature review. Only one study utilised a 'think-aloud' method in order to understand the limitations of using CA techniques in a context of patient-centred care (Cheraghi-Sohi et al. 2008). The limitations identified in this study seemingly impact the validity of the DCE method in dramatic ways, highlighting the influence of 'personal experience' upon the hypothetical nature of the task (Cheraghi-Sohi et al. 2008). Similarly, the selection of attribute levels was very superficially referenced and rarely justified.

Because attributes and levels give rise to numerous configurations all studies used a fractional-factorial design, using computer software to reduce the number of choices presented to respondents. This is done because presenting all possible combinations of attributes and options has proven to be unmanageable for respondents. Respondents can generally manage between 9 and 16 comparisons before becoming overburdened (Ryan et al. 2001). Overall, researchers using CA techniques attempt to reduce the cognitive burden of the tasks in order to avoid gaining inaccurate data.

Conjoint analysis studies typically have between 150 and 1200 respondents (Orme 2010). Final sample size ranged from 51 to 3,893 participants with a median of 666, suggesting robust studies overall. However, some studies used a sample size below 300 (with one study using less than 100). There is evidence to suggest that these studies would be underpowered. Four of the studies had seemingly inadequate (small) sample sizes (Vick & Scott 1998, Ryan et al. 1998, Markham et al. 1999, Haas 2005, Fiebig et al. 2009). The response rates (from 18% to 94%) were reasonable, despite not finding any evidence of direct motivation and encouragement to respondents.

It can be inferred that the CA rating and choice-based methods were not too demanding to complete in a pen-and-paper context, whether postal or recruited face to face. However, it is important to conduct developmental testing to develop the questionnaire to maximise potential response in the main study.

(c) The extent to which validity has been assessed in these studies

Studies reflect the scarcity of external validation studies of CA methods (Mark & Swait 2004). Another aspect of validity was tested in fourteen studies by comparing their findings with what might be predicted by existing theory. One study looked at test–retest reliability (Caldow et al. 2007). Only one study compared DCE estimates with a rating scale (Hjelmgren & Anell 2007), concluding that DCEs allow researchers to be more precise about the trade-offs, as opposed to direct ratings which allow all attributes to be considered equal.

Findings demonstrate that there is mixed opinion about the removal of irrational responses in analysis. The most commonly employed test for rationality was dominance testing (though even then some studies did not exclude these responses from analysis). Overall, this review reflects some emerging problems in rationality testing identified by researchers in the field i.e. the notion that irrationality can be due to limitations in design and implementation of DCEs, learning about their preferences or tasks and irrationality' tests not being conclusive (Lancsar & Louviere 2006). The included studies support the notion that deleting valid responses may lead to bias and lower statistical efficiency, and this is an important area to consider in the design of the main CA study in this thesis.

Use of the ISPOR checklist

The fair level of inter-rater agreement reflects a combination of weaknesses in the knowledge of the reviewers, in the ISPOR checklist itself and in the clarity of reporting in the reviewed studies.

Both reviewers were inexperienced in the field of CA study design. Subject-matter knowledge is important in the consistent appraisal of methodological quality.

Although it is made clear in the checklist outline that it should not be interpreted as endorsing any one particular CA approach or method, the clarity of the ISPOR checklist was questionable during practical application. Firstly, there are some compound items, e.g. in item 8.2: 'was the quality of responses examined, for example (rationality, validity, reliability)?' Secondly, some items contained ambiguity, e.g. '4.3 (should) an opt-out or a status-quo alternative (be) included?' It is uncertain how to mark these sections if only a part of such items have been satisfied. Many of the items depend upon completeness of reporting of the questionnaire and instructions to participants, though these details are seldom possible to present within the constraints of a journal article. This might be a case for publication of protocols and data collection instruments. Thirdly, there are some obscure terms that were hard to judge, for example '5.1 were the tasks 'sufficiently motivated?' This was hard to interpret and consensus was achieved only by setting some rules, for example a study was sufficiently motivated if there was a high response rate.

3.5.3 Evidence gaps in the included studies

Strengths of the included studies were that the majority gave information about attributes, used statistical analysis that was appropriate and were

comprehensively reported. Issues of validity and study limitations/generalisability were also discussed. The vast majority of studies used accepted methods of regression analysis like regression and probit, logit or mixed logit models, indicating a certain robustness of the included studies. However, the use of the ISPOR checklist as an accurate means of quality assessment remains unclear. Moreover, the majority of used a discrete choice format, one that appears to be a familiar behaviour for people. Multinomial logit analysis provides a well-developed statistical model for estimating respondent partworths from choice data. However, the choice format provides less information about the strength of preference or the order of importance of the rejected choice alternatives than ranking or rating. Making a choice only gives which alternative is preferred. For this reason the choice design requires larger sample sizes than ratings-based techniques (Johnson & Orme 2007).

However, there appeared to be several evidence gaps identified from the reviewed studies. The majority of studies lacked a detailed description of the trade-offs between the attributes. There was a distinct lack of qualitative research to understand decision-making behaviours, with only one study utilising a 'think-aloud' procedure in order to understand these potential limitations better (Cheraghi-Sohi et al. 2008).

Few studies compared DCE findings with other methods, such as a rating scale (Hjelmgren & Anell 2007). For example, two studies excluded in the review (for not meeting the requirements set out in the PICO inclusion criteria in Table 3.1 on page 58) found that there was discordance between DCE and direct rating exercises. However, it was unclear whether this was an artefact of different

designs, or respondents did not have clear preferences at the time of testing (Caruso et al. 2009, Hundley & Ryan 2004).

Other limitations of the included studies became apparent in the light of other studies that were uncovered but excluded. Although none of the reviewed studies included attributes based around the doctor as a whole person, i.e. their personal attitudes and beliefs (dimension 5 of the definition of patient-centred care in Table 3.2), three studies identified during screening used CA to investigate how the doctor as a person, i.e. their unconscious personal subjectivity (attitudes and beliefs) influenced the preferences of practitioners within rehabilitation counselling. One utilised a secondary care practitioner and two (using rating-based card sort exercises, consisting of 55 cards) used a trainee counsellor student sample (Wong et al. 2004, Rosenthal et al. 2006). Both concluded that personal prejudices and attitudes towards patient age and gender significantly influenced practitioners' preferences towards rehabilitation.

Although a small number of the included studies tested for respondent rationality and consistency and used these as criteria for exclusion, the studies highlight that respondent rationality and consistency is rather less straightforward to detail. CA may detect attitudes and preferences (and prejudices) that people may not necessarily admit to, be consciously aware of or be able to rationalise. CA could potentially be utilised for uncovering determinants of GP consultation in a similar way, for example, might unconscious factors like age or gender of GP influence a patients' decision to consult. Similarly, respondents may be learning about their preferences as they complete the CA tasks and thus may change their preferences from the beginning to the end of the questionnaire (rendering the tests for consistency unjustified). These studies reflect the request for increased

attention to the sense that removal of irrational and inconsistent responses means a removal of valid responses (Lancsar & Louviere 2006).

3.5.4 Strengths and limitations of this systematic review

A strength of this review was that an inclusive initial search strategy was used to locate relevant studies, reducing the chance that relevant studies related to CA patient-centred care in the GP consultation were missed. This process also allowed for a comprehensive collection of methodological studies and references as well as a comprehensive appraisal of previous published findings and reviews. Search terms were tailored to each database used. At the time of conducting the review (August 2009), the conjoint analysis terms project of the conjoint analysis working group of ISPOR had not yet been published. This may have provided an authoritative source for search terms (http://www.ispor.org/sigs/PRO_PPMI.asp).

A potential limitation of this review is that quality assessment was based on the recently developed ISPOR checklist – something designed as a guide for researchers designing CA studies and not primarily for reviewing them - and there was only fair agreement between the independent reviewers. In the main weaknesses of the knowledge of the reviewers (and weaknesses in the ISPOR checklist) appeared to influence levels of agreement.

Another limitation was that the definition of patient-centred care as inclusion criterion was hard to operationalise consistently. For example, including 'continuity of health professional' as the starting point for patient-centred care (i.e. patient-centred access to care) in this review may be controversial, because although Berry et al. (2003) defined 'continuity of care' as an aspect of 'patient-centred access to care' (Berry et al. 2003), Cheraghi-Sohi et al. (2008) defined 'continuity of care' as an aspect of relational continuity. This review included 'continuity of

care' on the basis that it is a patient-centred aspect of care and thus in terms of looking at attributes relevant to the decision to consult the GP for joint pain, it was relevant to the aims of the review.

The scope of the review included patients' preferences for general practice consultation contexts only and excluded studies into secondary and tertiary care settings. However, it is debatable whether some of the included studies may have belonged in secondary care, or at least a more specialised primary care consultation. For example, Fiebig et al's (2009) study into pap tests. The ever-blurring distinction between primary care and secondary care meant that Fiebig et al. (2009) study was included.

Similarly, locating studies within the context of GP consultation was problematic because the term GP is not necessarily applicable overseas, for example in Sweden and the U.S different terms are used to refer to the GP. Also in the U.S the healthcare system is organised differently so this may have further implications for the relevance of their studies in terms of this review.

3.5.5 Implications for future research

Overall, CA techniques offer the potential for a fuller description of the decision-making process of patients accessing primary care. However, three main considerations for future research have been highlighted by this review:

i. More researchers are beginning to claim that deleting irrational and inconsistent responses may lead to bias and lower statistical efficiency. For example, the latest software Adaptive Choice-based Conjoint (ACBC) was designed to adapt and enable respondents to learn about (and narrow down) their preferences as they go along. The notion that preferences do not necessarily pre-exist prior to the decision task and may change

throughout the task itself is one important area of future research (Ryan et al. 2001).

ii. Research into the usability of the ISPOR quality assessment checklist is needed.

iii. Although external validation studies were not searched for specifically, there is a lack of external validation of CA studies of patient-centred care in the GP consultation. CA is criticised for assuming a model of decision-making that lacks external validity and for being highly sensitive to issues of wording and context. In the context of patient-centred care attributes it is specifically an area that could benefit from external validation because it appears that these non-technical aspects of care are important to respondents, though they are hard to measure in terms of what actually happens in the consultation. The patient-centred attributes are in many ways centred upon the perception of the patient and it would be challenging to validate this externally.

3.5.6 Implications for thesis

This review has highlighted that there are some aspects that are of chief importance for this thesis:

i. One important decision in a CA study is which mode of administration to use. The majority of reviewed studies administered their survey in a postal format and all studies used a pen-and-paper format. However, computer-based CA is becoming more popular so it will be important to conduct developmental tests to investigate the feasibility and acceptability of the computer-based CA methods.

ii. It is important to consider the number of attributes that participants can be reasonably expected to manage. The number of attributes ranged from to 3 to 7. Most surveys included 5 or 6 attributes, and the majority of DCE's presented pairwise choice sets to each respondent. In this more personalised decision that involves complex factors it is important to undertake developmental studies to explore any cognitive limitations of people with osteoarthritis.

iii. This review highlights that only one study used a 'think aloud' in the development of their wording and design (see section 4.2 for 'think aloud' definition). In this study the 'think-aloud' procedure identified the influence of 'personal experience' upon the hypothetical nature of the task (Cheraghi-Sohi et al. 2008). Cognitive interviewing will be used to ensure that respondents understand the hypothetical nature of the task in the development of the design for this study.

iv. The choice of attributes and levels will be based on a review of the literature and qualitative developmental testing with the sample population. The attributes and levels used in the reviewed studies will also be considered in the choice of attributes and levels to use in the main study. For example, attributes such as thoroughness of technical care and shared decision-making attributes will be considered for inclusion in the developmental studies (see Table 4.4 on page 113). The specific wording used previously in the conjoint format may ensure comprehensibility of attributes.

v. In terms of determinants of consultation for OA this review has helped identify a range of attributes and levels (including the wording), which may be helpful for describing aspects of patient-centred care. This is relevant to

this thesis because there is evidence to suggest that aspects of patient-centred care (whether the GP takes the problem seriously or not) as well as clinical need related factors influence the decision to consult (Gignac et al. 2006).

vi. It is clear that within the reviewed studies, the choice of attributes have been limited to enabling factors ('patient-centred care', 'process' and 'interpersonal care') attributes. None of the studies have incorporated aspects of 'need related' attributes in conjunction with these attributes. Using these in conjunction may help to characterise the decision to consult the GP for joint pain more specifically, and provide an opportunity for embedding the Andersen-Newman model within a hypothetical framework which can quantify the relative importance of the key domains of healthcare use (including 'enabling' service configurations not currently available).

vii. Similarly, none of the studies included the GP's attitude towards their actual problem in conjunction with other patient-centred care attributes. Haas (2005) included 'doctor's recognition of pain', which is similar (Haas 2005). Developmental work will be crucial to testing the importance of this within the specific context of joint pain.

viii. It is unclear in these studies whether respondents were sufficiently briefed on how to complete the CA tasks. A demonstration exercise to explain the task may need to be given at the beginning, to bring motivation and encouragement to the task. This will be an area for developing and testing in preliminary designs.

ix. CA techniques are still being used in fairly unsophisticated ways in aspects of patient-centred care in the GP consultation. For example, using CA to

investigate the unconscious irrational factors that may influence decisions (as with the counselling rehabilitation practitioners who were influenced by the age and gender of their patients) shows that CA has the potential to get deeper into the unconscious and seemingly less rational factors influencing decision-making. CA may provide an opportunity to include unconscious factors (aside from clinical need related factors) that are influencing the decision to consult the GP, for example, the age and sex of the GP.

x. Only three studies excluded 'irrational' responses from analysis, in order to ensure that utility maximisation rules were consistent. The importance of including these will need to be weighed up while developing the final design.

xi. Existing research suggests that CA may have the potential to gain more accurate data about preferences than conventional direct rating exercises (Hundley & Ryan 2004). This systematic review identified a study in primary care services which reinforces this finding that DCEs allow researchers to be more precise about the trade-offs, as opposed to direct ratings which allow all attributes to be considered equal. (Hjelmgren & Anell 2007). This PhD thesis thus has the potential to build upon this existing work to reveal if there are differences between conventional rating methods and CA.

xii. Although the median response rate of studies was 60.5%, the range went from 18% to 94%, reinforcing the importance of developmental testing in order to ensure the CA task is acceptable and manageable for respondents.

3.6 Conclusion

Conjoint analysis estimates the relative importance of attributes when respondents are forced to choose between sets of choices. The included studies

claim that CA is preferable to other methods in this way. It appears that aspects of patient-centred care can successfully be quantified in the CA format, though no studies were identified that looked at clinical need related attributes in conjunction with patient-centred care attribute. There are limitations in these studies that will need to be taken into consideration in the design of the main study of this thesis. Firstly, there are a range of CA formats (the most popular being pairwise DCE in postal format) though it is unclear whether other formats or administration formats (like computer-based) were considered. Secondly, many of the studies derived attributes from the researcher's perception of what is important within the context, and failed to reach the gold standard hinted at in the ISPOR checklist (i.e. using a systematic literature review and qualitative methods). Finally, there is a distinct lack of qualitative research seeking to test the acceptability of the tasks to respondents. Further research into the acceptability of CA tasks within the context of GP consultation for joint pain is needed.

4 Chapter Four: Methods I: developmental studies and preliminary testing of respondent efficiency

4.1 Introduction

The primary aim of this thesis is to quantify the relative importance of selected clinical need and general practice service factors ('attributes') in the decision to consult the GP in older adults with joint pain/OA. Chapters 4 and 5 report on the methodological development work that has been undertaken in order to ensure the overall efficiency of the CA design.

The recommended stages and design principles set out in the ISPOR checklist (see Figure 3.1 on page 65) have been used to inform the definition of design efficiency in this chapter. The design efficiency consists of two aspects: respondent efficiency and statistical efficiency. Chapter 4 describes three developmental studies of respondent efficiency. This refers to the acceptability of the task to patients and is required in order to design an optimal conjoint questionnaire. Statistical efficiency, on the other hand, increases as the number of attributes and choice tasks presented increases (Cunningham et al. 2009). This is covered in chapter 5.

Chapter 4 therefore uses the qualitative methods recommended in the ISPOR checklist to gain the views of the Research Users' Group (RUG) based at the Arthritis research UK Primary Care Centre (ARUKPCC). The RUG is made up of people with experience of (and carers of close relatives) with painful long-term conditions. The RUG has 11 members and meets every four months as a forum. They are involved with activities such as giving advice on research design and reading and feeding back on research materials (such as questionnaires and cover and information letters to patients).

The first section of chapter 4 outlines the background for the need to undertake developmental studies (section 4.2). The following section outlines the overall aims and objectives of the developmental studies (section 4.3). The design, results and discussion of each of the developmental studies are covered in 4.4, 4.5 and 4.6. Section 4.7 presents a summary of findings and a final discussion is covered in section 4.8.

4.2 Background: the need for developmental studies

The selection of attributes and levels

The main objective of CA designs is to elicit respondents' preferences or priorities across the range of attribute levels that potentially characterise the service/product being tested. In this case it is important to identify the salient attributes, which characterise the decision to consult the GP for joint pain. Therefore, all potential attributes should be identified and considered (Bridges et al. 2011) before the salient ones are selected. It is crucial to describe all attributes and levels clearly and consistently to reduce the likelihood of respondents making unobservable assumptions about the attributes and levels presented in an alternative (scenario). Similarly the phrasing, description and supporting material presented to respondents for the attribute levels need testing in order to test whether or not they are understandable, plausible and whether respondents can undertake tasks effectively. For example, a previous CA study by Ratcliffe et al. (2004) suggested that pain (aches or pains) can be included as an attribute but it is important to test how respondents interpret this hypothetical attribute (i.e. pain) in a CA format in this context (Ratcliffe et al. 2004).

The construction of the conjoint tasks

As previously mentioned in chapter 2 (see section 2.2.3 on page 38) choice alternatives can include the full set of attributes included in the study (full-profile) or a subset of the attributes (partial-profile choice-based conjoint - PPCBC). According to Miller (1956) the capacity for processing information in individuals is limited to seven (+/- 2) pieces of information at one time (Miller 1956). PPCBC designs aim to reduce the cognitive burden for respondents according to this guideline. Thus prior to designing choice tasks with full-profiles, it is important to determine whether or not respondents can manage this task (i.e. looking at attributes salient to the decision to consult the GP) using full-profiles. It may be that the partial approach is better suited to this context.

Similarly, the number of choice alternatives included in each task will affect the respondents' cognitive processes. Moreover, the order in which attributes or questions are placed, or the number of attributes and levels must be tested to avoid potential measurement errors. The most important measurement error is the cognitive burden of respondents when too many attributes are presented to respondents at one time. Although, number of choice tasks increases statistical efficiency, it reduces respondent efficiency (it is too cognitively demanding), so it is important to minimise how many scenarios are included in each choice task – and how many attribute levels characterise them. The ISPOR checklist recommends a small cognitive debriefing interview and quantitative pilot in order to identify common misinterpretations of the instructions and whether the survey length is acceptable to respondents.

The data collection plan

CA questionnaires can be administered in multiple different ways, including interviewer-led pen-and-paper surveys, postal surveys (not interviewer-led), telephone or computer administrated on a limited number of sites, or web-based. Interviewer-led surveys may improve the quality of data because the interviewer can provide further explanation about the task where needed (Bridges et al. 2011). However, the presence of an interviewer during completion of the task may lead to more socially desirable answers being given by respondents (Bridges et al. 2011). It is important to design a method of data collection that motivates and informs respondents and to select an administration method that is most feasible for the target sample population (i.e. adults aged 50 and above). Moreover, in the instructions and layout respondents need to be informed, encouraged, and motivated (Bridges et al. 2011).

Recommended methods for developmental studies

The ISPOR checklist (Bridges et al. 2011) recommends four methods to assist the aforementioned developmental stages (i.e. the elicitation of salient attribute levels, the construction of the conjoint tasks and the data collection plan). These include the use of secondary research and data, cognitive interviews, researchers' knowledge and expertise and group discussions/focus groups. Each will now be presented in turn.

Use of secondary research and data

The identification of attributes and levels must be supported by secondary evidence obtained from existing studies (Lancsar & Louviere 2008). Sources of evidence can be quantitative or qualitative (secondary or primary) data. The most

important criterion is that the research is relevant to a) the potential range of preferences that people may have and b) includes individuals who represent the target population (Bridges et al. 2011). For example, Cheraghi-Sohi et al. (2007) constructed a 'map' of concepts relating to patient-centred primary care via a literature review (Cheraghi-Sohi et al. 2007).

Cognitive interviews

One emerging technique that has been used in CA studies is cognitive interviewing (CI) (Ryan et al. 2009, Cheraghi-Sohi et al. 2007). This was identified within the systematic review (see section 3.5.6 on page 94). CI techniques focus on the mental processes used by respondents to answer survey questions rather than the actual answers of respondents (Willis 1999). Drawing on information processing theory and research into memory recall, CIs are based on the premise that sources of response error relate to how respondents interpret and understand questions (comprehension), recall memories (recall), make decisions (judgment/decision-making) and construct answers (response) (Willis 1999).

Table 4.1: The four sources of response error according to cognitive theory

Four sources of response error	Examples of probes
1. Comprehension of the question	**Question intent:** What does the respondent believe the question to be asking? **Meaning of terms:** What do specific words and phrases in the question mean to the respondent?
2. Recall from memory of relevant information	**Recallability of information:** What types of information does the respondent need to recall in order to answer the question? **Recall strategy:** What types of strategies are used to retrieve information? For example, does the respondent tend to count events by recalling each one individually, or does he/she use an estimation strategy?
3. Decision processes/judgment	**Motivation:** Does the respondent devote sufficient mental effort to answer the question accurately and thoughtfully **Sensitivity/social desirability:** Does the respondent want to tell the truth? Does he/she say something that makes him/her look 'better'?
4. Response processes	**Mapping the response:** Can the respondent match his or her internally generated answer to the response categories given by the survey question?

By asking respondents to verbalise their internal thoughts, cognitive interviewers seek to reveal as much information as possible that may lead to insights about the respondents' decision-making processes. The core techniques used in CIs are 'think aloud', probing and observing (see Table 4.1). The 'think aloud' technique is a respondent-led technique, relying upon the respondent to think his/her thoughts aloud while they are thinking in response to a question. 'Think aloud' techniques have been endorsed (Ryan et al. 2009, Cheraghi-Sohi et al. 2007), suggesting that by asking individuals to verbalise their thoughts researchers can uncover the vital cognitive processes underlying the limitations of CA tasks. 'Think aloud' can be used during the interview (concurrent) or once the interview has finished (retrospective). The former is considered the most efficient, since thinking aloud while completing the cognitive task of answering the interview question removes recall biases (Willis 1999, Campanelli 1997).

To assist this process, probing techniques and observation skills are used in response to the respondents' verbalised thoughts with the aim of delving deeper into the their cognitive processes (see Table 4.2.)

Table 4.2: Some basic example of probes

Type of probe	Example
Specific probe	Why did you chose option A rather than option B?
General probes	Could you tell me more about your thinking on that?
Comprehension	What does this term mean to you?
Recall	What time period were you thinking of? (from when to when?)
Judgment	How sure are you that the doctor is giving a thorough examination?
Response	How easy or difficult did you find answering this question? Why was that?
Paraphrasing	Can you repeat the question in your own words?
Relevance of question	How relevant or irrelevant did you find that question/section?

Two recent studies using CI techniques alongside DCE's suggest that respondents' willingness to trade varies dramatically. Some respondents fail to grasp the hypothetical nature of the task, basing choices on their personal experience (Cheraghi-Sohi et al. 2008). When asked to imagine unfamiliar experiences (particularly health states) people tend to focus on the differences between those unfamiliar experiences and their personal experiences, and misconceive the importance of the unfamiliar experiences (Stiggelbout & de Vogel-Voogt 2008). There is also evidence to suggest that individuals are more likely to reinterpret attributes, relating them to other attributes outside the exercise, when considering concepts like quality of care, rather than with more tangible concepts like location (Cheraghi-Sohi et al. 2008, Gerard et al. 2008).

Researchers' knowledge and expertise

In addition to literature reviews and testing with a sample of the target population, it is suggested that clinical experts are consulted (Louviere et al. 2000). Discussions with experts can be useful for narrowing down extensive lists of attributes (Bridges et al. 2011).

Group discussions/ focus groups

The ISPOR checklist outlines that secondary data alone is insufficient to inform the development of the attribute levels and scenarios (Bridges et al. 2011). Popular methods of primary data collection include the use of discussion and focus groups. For example, Phillips et al. (2002a) used focus groups in order to explore the discordant valuations from a conjoint and traditional method (including likert rating scales). The decision to include discussion groups relies on various factors, including the nature of the questions asked and the respondents included

in the research. Basic thematic analysis is sufficient for identifying the total number of attributes and levels. Structured discussions with experts and further pilot testing with subjects are typically used to finalise the salient attributes (Bridges et al. 2011). However, discussion groups are susceptible to bias, which can negatively influence the development of the design. Miriampolski (2001) argues that disadvantages of focus groups include social desirability responses and the over reliance on an interviewer to stimulate debate and discussion (Mariampolski 2001).

4.3 Aims and objectives

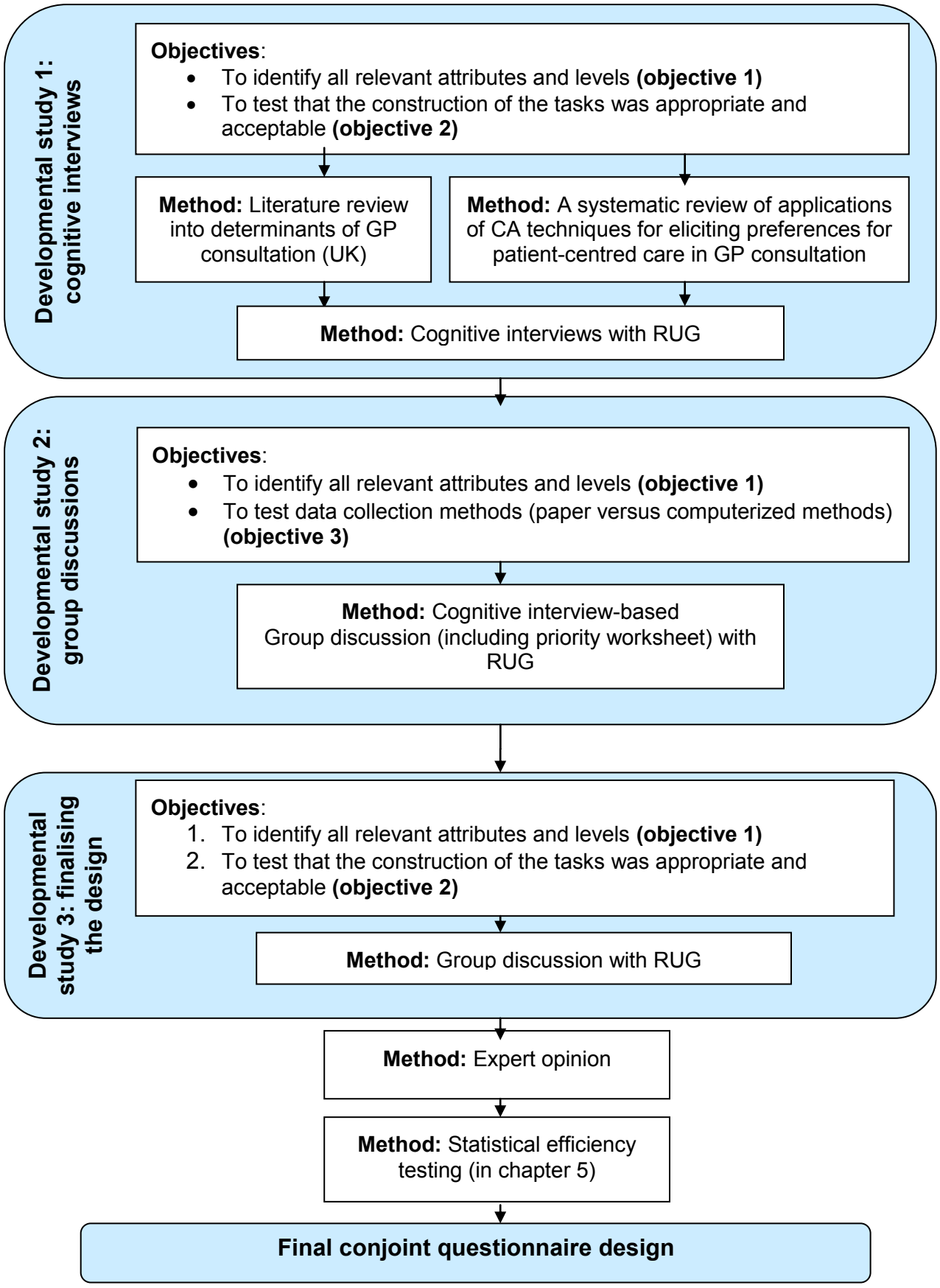
There were three developmental studies (1-3) (see Figure 4.1 overleaf). The aim was to ensure respondent efficiency in a sample of adults aged 50 and above. This is because the main study will target this age group.

1 To identify all relevant attributes and levels and to ensure these are supported by evidence from the literature and qualitative research (cognitive interviews/focus groups).

2 To test that the construction of the tasks was appropriate and acceptable (not overburdening) for respondents. The focus was on the amount of information presented within each choice task (and the plausibility of combining attribute levels together), the number of alternatives/profiles that could be feasibly compared within each choice task and the likely maximum number of choice tasks feasible for respondents to complete.

3 To test data collection methods (e.g. the instructions on how to complete the tasks and mode of administration). Most importantly are pen-and-paper or computer-based methods more appropriate for people in the target age range (50 and above)?

Figure 4.1: Flow diagram of developmental studies



The flow diagram in Figure 4.1 presents the objectives, methods and order of the three developmental studies that led to the final design.

4.4 Developmental study 1: cognitive interviews

4.4.1 Method

Design

The aims within developmental study 1 included: selection of salient attributes (to establish how relevant the attributes and levels identified in the literature search are and also identify other salient attributes), the construction of conjoint tasks and data collection plan (to get an idea of how feasible the CBC method was with people in the target age range - 50 and above).

The CBC task was developed based on the literature review into the determinants of consultation in older adults with joint pain undertaken in chapter 1 (see section 1.5.2 on page 12). The search terms used to identify relevant secondary research to inform the initial selection of attributes, levels and scenario's salient to the decision to consult the GP for joint pain (see Appendix 3a). For example, the term 'emotional functioning' was taken from patient reported outcome measures for pain in clinical trials (Turk et al. 2006). See Table 4.4 for full information on studies from where attribute levels were derived. Some of the wording of the attribute levels relevant to patient-centred care were taken from the articles identified in the systematic review into CA applications in chapter 3 (Markham et al. 1999, Haas 2005, Longo et al. 2006, Caldow et al. 2007, Cheraghi-Sohi et al. 2008) (See Table 4.4 on page 113).

The design, layout and wording of instructions for the first questionnaire tested were developed based on the DCE design developed in one study (Cheraghi-Sohi et al. 2008). This layout was chosen because it appeared to be the

most simple out of the reviewed studies. The questionnaire was then tested using 'think aloud' cognitive interviews.

A search into the applications of CA in osteoarthritis populations was also undertaken. This enabled us to identify what types of CA (and attribute levels) have been used in OA research as well as how acceptable the CA task is to respondents with OA. These studies may not have been identified in the systematic review in chapter 3 because studies were limited to patient-centred care attributes, so a separate search was undertaken.

Applications of CA have been emerging in the last decade within osteoarthritis research. Thirteen relevant studies were identified after searching five databases, but only 9 were primary studies. Out of these four looked at patients' preferences for characteristics associated with treatments for OA (Fraenkel et al. 2004a, Fraenkel et al. 2004b, Ratcliffe et al. 2004, Fraenkel & Fried 2008), but only one included attribute levels relevant to the CA design in this thesis. Ratcliffe et al. (2004) used 'joint pain' and 'joint aches' as single attributes in CA. Table 4.3 is from a study which aimed to investigate patient preferences for attributes associated with the efficacy and side-effects of treatment for OA (Ratcliffe et al. 2004). A discrete choice experiment was used.

Table 4.3: Attributes and levels used in relevant study

Attribute	Joint aches	Joint pains	Mobility	Risk of mild/moderate side-effects	Risk of serious side-effects
Levels	Very slight	Occasionally	Normal	1 in 4 chance (25%)	1 in 4 chance (25%)
	Moderate	2-3 times per week	Some difficulty	2 in 4 chance (50%)	2 in 4 chance (50%)
	Severe	Most days	Confined to chair	3 in 4 chance (75%)	3 in 4 chance (75%)

However, in this study there was little investigation into how easy or difficult respondents found considering 'joint pain' (and 'aches') as hypothetical attributes.

Previous research suggests that ageing may be associated with an increase in associative and automatic processes such as heuristics (Yates & Patalano 1999). However, it appears that these issues have not been explored in CA research to any great extent. There appears to be no existing studies using CA techniques to directly address issues of GP consultation within joint pain/ symptomatic OA. An understanding of the determinants of GP consultation using CA would provide an understanding of the relative importance of clinical need determinants and service 'enabling' factors. The most popular method of CA in OA research appears to be choice-based and although CA has proven easy to complete across a mixture of sociodemographic groups (Wittink & Bergestuen 2001) using ACA with older adults with OA has only been tested a limited number of times (Fraenkel et al. 2004a, Fraenkel et al. 2004b). Fraenkel et al. (2004a) investigated patient treatment preferences for knee OA, using computer-based face-to-face questionnaires with 100 patients with knee OA (achieving a response rate of 84%). Many older patients were willing to trade-off treatment effectiveness for a lower risk of adverse effects. Within the systematic review in chapter 3 it appears that choice-based methods with two service alternatives (pairwise comparison) per choice task are acceptable to people in the target age range (50 and above). Therefore, it was decided to test a pairwise choice-based conjoint method to begin with, and if this was acceptable then this would become the method used in the final design.

Participants

The sample consisted of three members of the Research Users' Group, all aged 50 years or over with joint pain. It is important that people in the target age range (50 and above) can cognitively comprehend and complete the CA tasks. Sampling was conducted around one characteristic: those respondents were above 50 years of age. Members of the Research Users' Group (RUG) agreed to be approached to participate in the design of research at the university. Members of the RUG above 50 years of age were contacted by post and three members agreed to participate in the cognitive interviews.

Data collection procedures

Cognitive interviews were used to investigate potential respondents' reaction to and understanding of a CBC task about the decision to consult a GP for joint pain. A CI script was developed and consisted of three sections; an introduction, ending and a section including general probes (see Appendix 3b). Initially the interviewer stressed to participants that he/she is not primarily interested in the data generated from the questionnaire, instead it is the questionnaire itself that is being tested. Before it is used widely it is important to assess potential participants' views on how to complete it and what the questions mean.

To reduce researcher bias one standard interview script was designed for all cognitive interviews. The general section of the script included notes to the interviewer like 'remember to reward respondents'. A combination of approaches was used, including concurrent 'think aloud', pre-planned (see Appendix 3c) and general probing (see Appendix 3d).

Two CIs were conducted: one included two respondents (a married couple) and another respondent was interviewed one-to-one. The 'think aloud' CI's were recorded for transcription.

The CBC exercise in in the cognitive interviews involved three sections. Section one was a direct rating task of seven single attributes (from 1-10, with 10 being most important) for the decision to consult the GP for joint pain. Section two included the CBC. This was based on four attributes with three levels each (3*3*3*3=81 profiles) drawn from the literature identified. As is conventional in CA (see section 2.2.1 on page 28) the combination of potential profiles was reduced using orthogonal arrays and thus reduced the burden for respondents (and still ensured an acceptable level of statistical efficiency). For further information on statistical efficiency see chapter 5). Attributes used included two clinical need factors (severity of pain and comorbidity), one on potential for curability and the fourth attribute was related to interpersonal manner of doctor (see Table 4.4 on page 113). As 'doctor listens' emerged from the systematic review as the most important attribute across studies, it was considered in this early stage of development. However, the GP manner seemed to be more comprehensive and relevant and so was included instead of 'doctor listens'. The CBC exercise involved five choice sets, and each choice set involved respondents' choosing between two scenarios as to which would most likely lead them to consult a GP for joint pain. Section three included respondent demographics (Appendix 3e).

The CBC focused on how respondents directly rated determinants of consultation for joint pain and how respondents made choices between different levels in a hypothetical CBC.

Table 4.4: attribute levels used in Developmental study 1

Attributes		Levels		Basis	Andersen-Newman (1995)
1	Pain Severity	1.1.	I have mild pain	(Sanders et al. 2004, Ratcliffe et al. 2004, Bedson et al. 2007, Rosemann et al. 2007)	Need factor
		1.2.	I have moderate pain		
		1.3	I have severe pain		
2	Other illnesses (comorbidity)	2.1.	I have no other illnesses	(Bedson et al. 2007, Kadam & Croft 2007)	Need factor
		2.2.	I have one other illness (e.g. hypertension/diabetes)		
		2.3.	I have more than one other illnesses (e.g. hypertension/diabetes)		
3	Curability	3.1.	The doctor can help to increase physical functioning	(Turk et al. 2006)	Enabling factor
		3.2.	The doctor can help to reduce my pain (severity, character and frequency)		
		3.3.	The doctor can help to increase my emotional functioning		
4	Doctor's access to medical notes and knows you	4.1.	The doctor has access to my medical notes but does not know me	(Haas 2005, Cheraghi-Sohi et al. 2008)	Enabling factor
		4.2.	The doctor has access to my medical notes and knows me a little		
		4.3.	The doctor has access to my medical notes and knows me well		
5	Thoroughness of physical examination	Not included in CBC exercise, only in rating exercise 1		(Cheraghi-Sohi et al. 2008, Markham et al. 1999)	Enabling factor
6	Doctor's interest in your ideas about problem	Not included in CBC exercise, only in rating exercise 1		(Cheraghi-Sohi et al. 2008, Scott et al. 2003)	Enabling factor
7	Time to get appointment	Not included in CBC exercise, only in rating exercise 1		(Cheraghi-Sohi et al. 2008, Caldow et al. 2007)	Enabling factor

After respondents had agreed to participate CIs were conducted on the university premises, participants were informed in a postal information sheet about the nature and type of questions that they might be asked (see Appendix 3f).

Respondents were asked to complete the tasks (see Appendix 3e) in the presence of the researcher. When there was more than one respondent both respondents were asked to take a turn in thinking aloud. At the beginning of the interview, instead of explaining the technical principles of the CBC, a familiar format (i.e. choosing a package holiday) was relayed to respondents to orient them to the task.

It is also important for the interviewer to train the respondent in performing a 'think aloud' procedure at the beginning of the interview. Respondents were asked to complete the 'windows exercise' in order to warm them up for the 'think aloud' process.

'Try to visualise the place where you live and think about how many windows there are in that place. As you count up the windows, tell me what you are seeing and thinking about' (Campanelli 1997).

The warm-up 'windows exercise' was considered important for preparing respondents to 'think aloud' and to demonstrate what type of data was required. The pre-planned probes were minimal and consisted of general reflective questions at the end of the interview.

If the respondent lost focus during the interviews 'think aloud' techniques were used by the interviewer to encourage the cognitive process. Similarly if a respondent made an unclear statement probing was used to delve deeper into the decision-making process. These probes were spontaneous and kept to a minimum to ensure the natural cognitive flow was not disrupted.

At the end of the questionnaire, respondents were asked pre-planned reflective questions about their feelings about the overall experience, the overall ease/difficulty of the task and the length of the questionnaire. The sessions were tape-recorded and transcribed verbatim. Respondents were offered an opportunity for expenses to be paid in full.

Analysis

The data was transcribed and examined to identify reoccurring and dominant themes across interviews. A basic thematic analysis was conducted. The common themes were then mapped according to the aims of the developmental studies set out in section 4.3.

4.4.2 Results of developmental study 1

This section presents the results of developmental study 1. An example of a full transcript can be found in Appendix 5d. Table 4.5 on the following page presents the results from the cognitive interviews.

Table 4.5: Results of developmental study 1.

Objectives (1-3)	'Think aloud' response errors	Transcript quotations	Potential revisions
1. Attributes and levels	Respondents agreed all included attributes were relevant, but suggested four additional attributes: 'Influence of others' (including spouse, friends and family), 'Self-medication or alternative therapies' 'Attitude of receptionist' and 'Information from the media' (TV/internet)	*'See one of the first barriers will be the receptionist; getting through' (male, 1)* *'I think some people get the feeling you know that they're there as a filter. Erm, to stop people' (male, 1)*	Further testing required
1. Attributes and levels	The term 'emotional functioning' was too clinical, unfamiliar and awkward for respondents and this caused problems. One respondent displayed a strong emotional reaction to the term 'emotional functioning'	*'The doctor can help to increase my emotional functioning. That's going to throw a lot of people. What's that?' (male, 1)* *'I don't like the term (emotional functioning). It brings back memories of my mother, my mother use to say to me well you've got to go; you can't go to school until you've been' (female, 1)*	Change emotional functioning to 'How I can deal with the problem?' (male, 81) A description of terms at introduction is needed
1. Attributes and levels	Respondents avoided 'thorough examination' assuming that they weren't qualified to accurately judge the GP's thoroughness based on their understanding of the term	*'Thoroughness of physical examination. Are we qualified to say whether it was a thorough examination or not?' (male, 2)* *'Can we really be sure he is being technically thorough? (male, 2).*	Introduce attributes at beginning. 'As thorough as you would like'

Table 4.5 cont: Results of developmental study 1.

Objectives (1-3)	'Think aloud' response errors	Transcript quotations	Potential revisions
1. Attributes and levels	The attribute 'doctor knows you a little' overlapped with 'potential for curability'. When both choices had the same level of 'doctor can help to reduce my pain' the 'doctors knows you' level swayed one respondent, thus demonstrating how clear attributes and levels are important	*'If it's severe, you've really got to choose then who's the better doctor? ' (female, 80)* *'Because some doctors are better than others especially if they've suffered themselves' (female, 80)* *'I'd still go for B. 'doctor knows you a little' (male, 1)*	Conduct focus group and decide on attributes for main study design, ensuring against artificiality of attributes and correlation
2.Construction of conjoint tasks	The effects of personal experience were apparent where one respondent refused to answer three choice sets due to them not matching his situation accurately	*'I don't want to choose either because they don't match my situation properly' (male, 2)*	Repeat 'hypothetical exercise' throughout. Use opt-out option (where a respondent can choose neither) as in ACBC
3. Data collection plan	The decision to consult the GP is a familiar task for patients and all respondents failed to grasp the hypothetical nature of the task. Respondents superimposed their personal situation (of pain and perceptions of health professionals) onto attributes	*'No, it's moderate really but I think that one applies more to me (female, 1)* *'I mean she does know me; she just tries to do her best for me when I can get an appointment' (female, 1)*	'Pain severity' must be considered in relative and not absolute terms
3. Data collection plan	Some respondents overlooked the instruction that it is only consulting the GP and not any other health professional	*'Oh no, it's got to be your GP hasn't it?' (female, 1)*	Embolden GP and repeat only GP throughout

4.4.3 Discussion

Overall, the 'think aloud' technique proved useful for developmental study 1 in terms of comprehension, recall, judgment and response. These four sources of response error contribute towards the 'acceptability' of the task to respondents. Results support findings from previous studies (Cheraghi-Sohi et al. 2007) and illuminate new issues. For example, influences of willingness to trade upon the validity of CBC's are well researched. However, current research examining the impact of emotions upon the decision-making process is limited. Potential revisions to the CBC were suggested as a result of the CIs (see Table 4.5 on page 116).

The selection of attributes and levels

CIs helped identify a relevant and plausible choice of attributes characterising the decision by older adults to raise their joint pain within general practice – something crucial to ensuring the validity of CBCs. Respondents agreed that seven of the attributes in the first direct rating exercise were relevant but all respondents suggested at least one other attribute that was important to them. These additional attributes were therefore considered for the design within developmental study 2 (including influence of spouse, friends, family and media, self-medication or alternative therapies and attitude of receptionist). A wide range of attributes emerged and this highlighted the importance of conducting a group discussion to assist in selecting salient attributes and levels. The influence of media messages is familiar in illness perception frameworks (Hale et al. 2007). However, the 'attitude of the receptionist' emerged as an unfamiliar attribute that may influence the decision to consult a GP for joint pain. 'Pain severity' is an absolute term and this caused problems in assisting respondents to get into the

'hypothetical' mindset for the task. A relative use of pain, for example 'pain gets worse' rather than 'pain is severe' might be more appropriate but would need more testing. Similarly, the term 'emotional functioning' taken from patient-reported outcome measures for pain in clinical trials (Turk et al. 2006) was too clinical, unfamiliar and awkward and would need rephrasing and testing in the next phase.

The results add to an understanding of the effect of emotional concerns upon CBCs. One study indicates that health policy issues can elicit emotional concerns among individuals, which might lead them to violate the compensatory rule (i.e. the value of good characteristics in products compensate for bad characteristics) (Arana et al. 2008). The effects of emotional processing need further consideration. Two respondents had problems judging the term 'thorough examination' assuming that they were not qualified to accurately judge the GP's thoroughness based on their understanding of this term. This may explain why it has been considered useful to fully explain terms in DCEs (Ryan et al. 2009). It is important to give a definition of the factors in the design at the beginning of the exercise and this needed developing.

There was evidence of a 'halo effect' of one attribute over another. A halo effect is '*when evaluations of one attribute spill over to evaluations of other attributes'* (Phillips et al. 2002a, p. 1667). This seemed to affect one respondent. The 'doctor has access to my medical notes and knows me a little' overlapped with 'the doctor can help to reduce my pain', where the latter was dependent on the former in the respondents mind to the point that the latter became overlooked and considered artificial. It is crucial that there is little overlap between attributes and levels because any overlap between attributes renders them insignificant. This finding demonstrates how it is crucial to clearly define attribute levels, especially

with complex concepts like patient-centred care (Cheraghi-Sohi et al. 2008). The 'doctor can help' levels may be replaced with a more objective measure of how much the doctor can help, for example, 'the doctor can offer you effective treatments', thus preventing it overlapping with 'whether the doctor knows me a little'.

The construction of the conjoint tasks

All respondents failed to understand and comply with the hypothetical nature of the CBC task, superimposing their personal situation and experience onto the attributes instead. Results reflect existing research to suggest that patients 'go beyond' the attributes given, incorporating the wider context of their lives (Cheraghi-Sohi et al. 2007). It was these formal constraints that 'bemused' and irritated one respondent to the point where he refused to complete the task. This was an important early finding in the development of the main study protocol because it outlined how crucial it is for respondents to understand what they are being asked to do and thus avoid providing inappropriate and invalid answers. More specifically it outlined the importance of using relative terms in an attempt to engage respondents so that they could respond based on their experience, thus limiting the chances of them refusing to participate (Phillips et al. 2002a). Failure to comply with the hypothetical exercise was the same for both clinical need attributes (i.e. 'pain severity' and 'other health problems') and perceptions of the health professional ('doctor can help me'), suggesting that these highly personalised states might be better presented in relative terms.

DCEs depend on a willingness and capability to trade and there was evidence of trading in some cases and not in others. In the cases where there was trading this usually involved trading between a selected few attributes already

deemed most important. One respondent clearly demonstrated that the severity of pain was the most salient attribute and remained consistent throughout.

One respondent changed their priorities at different choice sets. The respondent originally identified 'doctor knowing you' as more important than 'pain severity' and then later prioritised 'pain severity', though this seemed more indicative of the presence of her spouse, where it seemed they influenced each other's choices dramatically. Having two people in a CI proved challenging (especially two people who knew each other well) because there was collaboration in giving responses and led to preferences being changed. One respondent went so far as to refuse to finish the task based on his unwillingness to trade, due to the choices not matching his preference perfectly.

The data collection plan

The failure to be hypothetical is a significant limitation because in the context of large scale postal distribution, response is essential for valid survey results. For example, missing data within returned questionnaires can result in loss of precision, i.e. reducing the representativeness of the sample and leading to incorrect assumptions being made about the population being studied (Bowling, 2005). It is more desirable to attempt to reduce the possibility of missing data during the design of the questionnaire.

Previous research suggests that both response and completion rates are significantly affected by the 'acceptability' of the task to respondents and perceived difficulty (Ryan et al. 2001). Some of the issues of acceptability have been detected in this developmental study. For example, the specific instructions for the hypothetical task and use of language that is accessible to the lay reader are important for eliciting valid information. This is required in order to make it clear

that this exercise only relates to consulting the GP and not any other health professional. The instructions and layout need to be amended.

Studies that use computationally advanced methods, like adaptive choice-based conjoint (ACBC) that provide an opt-out option (where a respondent can choose neither and opt-out of the choice task), might also be appropriate.

4.4.4 Implications for main study design and further developmental work

Decision-making is a complex phenomenon and at each level a variety of cognitive (and emotional) factors can influence older adults in their decision to raise their joint pain within general practice. Many studies using DCEs fail to investigate these underlying limitations in the development stages of the DCE task. Cognitive interviews are one method for better understanding these limitations and were useful in this phase of developmental testing for this thesis. The following issues were highlighted and taken forward into developmental study 2.

- Uncertainty remains on which attributes to include in the main study.
- Attribute levels must be defined succinctly (with no artificial boundaries or halo effects) in order to allow accurate judgment and trading-off. The importance of using familiar, non-clinical language was highlighted. A description of the attributes and attribute levels in the introduction is required. Attributes must be described in a way that respondents can understand and judge easily.
- 'Pain severity' and 'other illnesses' must be considered in relative and not absolute terms. The perceptions of the health professionals may not be

included as an attribute but as subgroup analysis (i.e. sociodemographics) criteria instead.

- Respondents' interpretation of the task is mixed; some see the task as a preference rather than a matching exercise for their own personal experiences. A question remains on how to outline the hypothetical nature of the task in a CBC design for postal administration.
- Despite two respondents claiming it was 'easy' difficulties emerged within the interviews and one respondent resisted the choice task section completely. The option to include an opt-out may be an important feature for engaging and sustaining the willingness to trade for respondents and requires further work.
- A demonstration exercise may be required at the start of the questionnaire in order to introduce and warm up respondents for the choice task.
- The feasibility of DCEs for the over 50 population requires more investigation (with a larger sample size than this study).

Therefore, further testing is required.

4.5 Developmental study 2: group discussions

4.5.1 Method

Design

Developmental study 2 involved testing an adaptive choice-based conjoint (ACBC) design, an emerging computerised method of CA (see section 2.2.3 on page 38 for more information). Developmental study 2 aimed to specifically test the relevance of the attributes. It was also important to test the feasiblility of a full-profile ACBC computerised method with the target age group (adults aged 50 and above), by observing their decision-making heuristics using 'think aloud'. If ACBC

were to be used for the main study it would have to be conducted at the university premises (due to limited software availability). Thus, it was important to test whether respondents would be comfortable coming into the university and using the computer to complete the ACBC task.

The attributes used were developed based on the literature review in chapter one and involved a 'think aloud' cognitive interview (see section 4.2.4) and a discussion group centred around the cognitive interview.

The full-profile ACBC task was based on the review of the literature for determinants of consultation in older adults with joint pain (in chapter one) and the results from developmental study 1. The CBC was based on six attributes with 15 levels (3*2*3*2*3*2=216 profiles), but the combination of potential profiles was reduced in order to reduce burden for respondents (and still ensure statistical efficiency). For further information on statistical efficiency see chapter 5). Two attributes focused on clinical need factors (pain compared to usual and likely pain relief) and four focused on enabling attributes; GP's manner, GP can offer you, amount of information provided and doctor you see (see Table 4.6 on page 127).

The questionnaire was constructed manually on a computer by the researcher using the ACBC demonstration software provided as a free demonstration by Sawtooth Software (Inc. Orem, UT). The questions for each section were based on the template questions provided in the software and these were adapted to be appropriate to the decision to consult the GP for joint pain. The questionnaire involved three sections (see Appendix 4a).

i. A 'build your own' section introduces attributes and levels and asks respondents to indicate their preferred level for each attribute.

ii. Screening questions, profiles (scenarios) are shown four at a time per screen, for 7 screens. This section does not ask the respondent to make a final choice, but asks them to consider each one 'a possibility' or 'not.'

iii. The choice sets (three presented at a time), involved respondents choosing between two scenarios as to which would most likely lead them to consult a GP for joint pain.

As part of this study respondents were also asked to fill in a 'priority worksheet exercise' at the end of the session. This gave an overview of the attributes used and asked respondents to prioritise 6 salient attributes and play the role of the researcher in deciding on them (see Appendix 4b).

Participants

Members of the ARUKPCC's Research User Group were invited to participate in developmental study 2, by postal invitation. There were 4 females and 6 males (all aged 50 years or over). None of these participants had been involved in developmental study 1.

Data collection procedures

The CI script used in this test consisted of three sections; an introduction, an ending and a section including general probes. The warm-up 'windows exercise' was used again, as in the previous test, to demonstrate what was wanted of respondents. To reduce researcher bias one standard interview script was also used (Appendix 3b). A combination of approaches was used, including concurrent 'think aloud', spontaneous and pre-planned probing (see Appendix 4c). The questions were very broad, for example, what did you think of that? and how easy or difficult was that to do?

The ACBC task focused on how respondents compared a series of different scenarios (referred to as situations in the task for a lay audience) and choose which one they thought would more likely lead them to consult the GP for their joint pain. However, ACBC (unlike the paper-based CBC in the previous test) adapts to respondents answers and is administered via computer.

After respondents had agreed to participate one cognitive group interview was conducted on the university premises. Participants were informed in a postal information sheet about the type of questions that they might be asked (see Appendix 4d).

One respondent was asked to complete the tasks (see Appendix 4a) in the presence of the researcher and members of the discussion group. One 'think aloud' was conducted based on an overhead projection of an ACBC questionnaire on a computer screen with a group of 10 members of the RUG. After each ACBC task was completed, a five to ten minute group discussion was instigated by researchers (DC, CJ, MF, BA-O), allowing issues that arose during the previous screens to be discussed. The session was tape-recorded and transcribed verbatim. Respondents were offered an opportunity for expenses to be paid in full.

Table 4.6: Attributes and levels used in developmental study 2

Attributes		Levels		Basis	Andersen-Newman (1995) model categorisation
1	Pain compared to usual	1.1.	The pain I am experiencing is less than usual	(Ratcliffe et al. 2004, Sanders et al. 2002, Hagen et al. 2000, Hawker et al. 2008)	Need factor
		1.2.	The pain I am experiencing is the same as usual		
		1.3	The pain I am experiencing is more than usual		
2	Likely pain relief	2.1.	15 % (Minimal)	(Caldow et al. 2007, Turk et al. 2006, Hawker et al. 2008)	Need factor
		2.2.	30% (Moderate)		
		2.3.	50% (substantial)		
3	GP's manner	3.1.	The GP will be friendly and Informal	(Donovan & Blake 2000, Morgan et al. 2000, Rosemann et al. 2006)	Enabling factor
		3.2.	The GP will be businesslike		
4	GP can offer you	4.1.	Painkillers and physiotherapy	(Jinks et al. 2007, Sanders et al. 2002, Tallon et al. 2000)	Enabling factor
		4.2.	Painkillers only		
5	Amount of information	4.1.	None	(Rosemann et al. 2006, Mallen & Peat 2009, Mann & Gooberman-Hill 2011)	Enabling factor
		4.2.	A little		
		4.3.	A lot		
6	You always see	4.1.	The same GP	(Caldow et al. 2007)	Enabling factor
		4.2.	Different GP		

Analysis

As with developmental study 1, the data was transcribed and examined to identify reoccurring and dominant themes across interviews. A basic thematic analysis was conducted. Data was then analysed according to the aims of the developmental studies set out in section 4.3 on page 109.

4.5.2 Results of developmental study 2

An example of a full transcript can be found in Appendix 5d. The main results from study 2 are found in Table 4.7 on the following page.

Table 4.7: Results of developmental study 2

Objectives (1-3)	'Think aloud'/Discussion group findings	Transcript quotations	Potential revisions
1. Attributes and levels	Painkillers and physio are already on the list for most people with OA, so alone these two are a very obvious choice. Some additional attributes were suggested	*'I did find it difficult that you had tied it down to painkillers and physiotherapy … because when I use the GP I usually hear of something else like the 'pain clinic' and I want to find out if the doctor will send me to the pain clinic or if the doctor will send me to have X ray's, so couldn't there be a third thing which gives a choice of 'other' or further treatments' (Respondent female 4)* *'I would want to see something like 'a new form of treatment'. (Respondent female 4)* *'Other treatment; holistic care, acupuncture, swimming, X ray' (Respondent female 4)*	Identify NICE guidelines for treatments and consult a GP/expert researcher: Painkillers only A promising 'new' treatment Referral for an X-ray
1. Attributes and levels	In the 'priority worksheet exercise' a range of attributes were suggested	*'The distance to your GP, psychological effects, ability to explain symptoms, help with sleeping because of pain, time to see GP from (different or same), waiting time for treatment/length of time before painkillers are effective, other health problems', race /age/gender of GP, reassurance from/approachability of GP'*	Other health problems' and ' GP's attitude/manner/appr-oachability' to be considered for further testing
1. Attributes and levels	The effects of the pain are not acknowledged in the existing attribute. Incorporating disability/ mobility and ability to work and perform certain tasks needs inclusion to ensure clarity	*'I think it depends if it's affecting your life, I mean if you're working and you're not able to go to work' (Respondent female 2)* *'Well, It's affecting your way of life isn't it?' (Respondent female 2)* *'Pain affecting work or everyday living'(Respondent female 4)* *'Condition is affecting my day to day life'(Respondent male 4)*	Change to 'disruption of pain to your everyday life compared to usual'. Not disrupting/Some/Most of everyday life. Describe attributes and levels at beginning to orient respondent

Table 4.7 cont: Results of developmental study 2

Objectives (1-3)	'Think aloud'/Discussion group findings	Transcript quotations	Potential revisions
1. Attributes and levels	Respondents agreed that all of the attributes in the ACBC exercise were relevant. However, many others were suggested. Only 6/10 of the worksheets were fully completed. 6 people prioritised different attributes, though many of them included the attributes already included in the questionnaire	'*They are all relevant really?' (Respondent female 2)* • *Pain* • *Disability/Independence/ Affecting work or everyday living* • *Known GP* • *Degree of likely pain relief* • *Age/gender/race of GP* • *Information available* • *Pain compared to usual level/ Present level of pain* • *Likely pain relief* • *GP can offer you* • *Amount of information* • *How would I like my GP to be* • *What would I like my GP to offer me* • *How much is my condition affecting my day to day life* • *Treatment my doctor offers* • *How much information will the GP give me* • *GP manner* • *Painkillers and/or physio* • *Amount of information* • *Distance of GP from home* • *To permanently judge pain on daily basis* • *Understanding and advice from GP* • *Imperative for full understanding* • *Consistency of opinions.* • *Expectation of painkillers and we have to live with it* • *General conversation with GP to X ray or acupuncture* • *Positive help*	Keep the six attributes included as they appear salient to RUG and make any necessary adjustments to levels

Table 4.7 cont: Results of developmental study 2

Objectives (1-3)	'Think aloud'/Discussion group findings	Transcript quotations	Potential revisions
1. Attributes and levels	The 'likely pain relief' needs clarifying in terms of whether it is what patient wants, expects or what the doctor can offer. The 15%, 30% and 50% was considered clinical, unfamiliar and meaningless	*'This likely pain relief, is it what you want, what you expect or what you actually think your doctor can manage to give you' (Respondent male 2)* *'Take out the percentages and just leave minimal, moderate or substantial, just use the words. I think the percentages might confuse people' (Respondent male 2)*	Change to treatments and services (based on NICE guidelines) rather than pain relief
1. Attributes and levels	GP's manner levels were considered artificial	*'I'm concerned about this false dichotomy between friendly and businesslike and informal and businesslike cause I think that is entirely false' (Respondent male 5)*	Combine approachability and GP manner: test GP's attitude
1. Attributes and levels	The influence of illness perceptions (i.e. cause and curability) of OA and perceptions of health professionals upon the decision to consult were suggested	*'Attitude of patient to treatment', 'Fear of what will happen in the future', 'Because of age, is it arthritis' 'Information on illness' ,'pointed in the right direction'*	Respondents to be subgrouped on the basis of illness perceptions questionnaire (IPQ-R)
2. Construction of conjoint tasks	There is too much information on the screen at one time for respondents to manage in terms of memory	*'The difficulty is keeping every one of those six factors in your head at the same time when also you've been given three different sets or circumstances, so for example you are looking at less pain and you think to yourself well you I wouldn't go to the doctor if I had less pain and then you say well there's a lot of information down here and you see a different GP and you go to that one, forgetting that you've decided beforehand that you wouldn't go down that column' (Respondent male 3)*	Present 2 scenarios rather than 3, ensuring each is labelled. ACBC design could be designed using only 2 options at a time, though labelling is not possible. A pen-and-paper-based CBC will be considered

Table 4.7 cont: Results of developmental study 2.

Objectives (1-3)	'Think aloud'/Discussion group findings	Transcript quotations	Potential revisions
2. Construction of conjoint tasks	One respondent felt that if profiles were on separate pages it would be clearer	*'If it is possible to separate them (the profiles) instead of them all being on the page, it would just delineate them even more clearly' (Respondent male, 2)*	Consider 2 profiles for comparison
2. Construction of conjoint tasks	It is unclear whether the attribute 'pain compared to usual' can be combined with 'likelihood of pain relief'	*'I do see some dichotomy between the top line 'pain is usual' and the likely 'pain relief' you will get' (Respondent male 2)*	It is unclear if these attributes can be plausibly combined
2. Construction of conjoint tasks	One respondent demonstrated a trade-off apparent with the pain severity and GP continuity. Another demonstrated that the severity of pain was the most salient attribute. Another respondent demonstrated that there was a trade-off against pain, based on an underlying perception that the health professional cannot help	*'I don't want to see Dr X, but if its urgent I must go in, but on the other hand only Dr X knows my circumstances' (Respondent male 3)* *'If you're GP wasn't friendly and you were in pain you would still have to go' (Respondent female 2)* *'Being able to get help from the pain clinic, I would put that before pain because pain to a doctor doesn't seem to, well I wouldn't say mean anything, but there's not much more they can do' (Respondent female 3)*	Respondents are clearly starting to make trade-offs. Unsure whether it is necessary to include opt-out profiles, as this may allow them to refrain from trading off and gaining an understanding of their priorities
2. Construction of conjoint tasks	Respondents were confused and frustrated by the similarity of choices/questions	*'You look at it and think I've just filled that in, why have I got to fill it in again? You're liable to skip it' (Respondent male 1)* *'You could put 'here are 4 more scenarios' (Respondent male 3)*	WRITE 'here are TWO more sets of situations that may lead you to consult the GP'
2. Construction of conjoint tasks	The 'think aloud' respondent failed to grasp the hypothetical nature of the task. Instead, his personal experience superimposed previous consultation experience onto the tasks	*'Are we not actually putting in the actual experiences we had when we last visited the GP? You cannot fill in for your next visit really because you don't know, all you can go by is your last visit" (Respondent male 1)* *'There is not really a choice for me there? I don't expect the pain to get any less, so I chose 15% per cent (Respondent male 1)*	Embolden and repeat. PLEASE REMEMBER THAT CHOICES ARE HYPOTHETICAL

Table 4.7 cont: Results of developmental study 2.

Objectives (1-3)	'Think aloud'/Discussion group findings	Transcript quotations	Potential revisions
2. Construction of conjoint tasks	It is unclear whether the attribute 'Pain compared to usual' can be combined with 'likelihood of pain relief', in terms of what is driving the person to go and what can be expected on arrival	*'I do see some dichotomy between the top line 'pain is usual' and the likely 'pain relief' you will get' (Respondent male 2)*	It is uncertain as to whether these attributes can be used as discrete choices together
3. Data collection plan	All respondents disapproved of a web based/computer option. Respondents felt the computer option limited their ability to go back and change answers	*'This is why I was thinking that people might want to go back, if there is a way to go back and change, there is no option to go back and change' (Respondent male 1)* *'No, they would much rather have a piece of paper where they can sit and quietly fill it in at their leisure' (Respondent female 1)*	Re-consider a postal pen-and-paper-based CBC design
3. Data collection plan	The central question 'what will encourage you to consult the GP' is ambiguous and needs rewording in simple language and repeated consistently at every page	*'What really needs making plain is: what will encourage you to consult the GP. Emphasise that' (Respondent male 1)* *'Remember will I go to the GP because of this' (Respondent male 1)*	Reword the central question. Include a 'demonstration of how to fill out the choice tasks' at the beginning and repeat 'this is a hypothetical task' throughout
3. Data collection plan	'You always see' was considered artificial and unrealistic	*'I think on the last one instead of 'You always see' I think if you put 'You can always see' that would be clearer. I would think 'Can I always see', 'yes you can'. And if I can't I just tick different GP (Respondent male 3)*	Change to: 'You can see'
3. Data collection plan	Respondents were confused by the layout and asked for questionnaire to be labelled and clearer	*'I would want each of those columns slightly separated from each other and then you can put here are four scenarios that may encourage you and then itemise it at the top of each one, scenario 1, 2,3 and 4 and that would make it clearer' (Respondent male 2)*	ACBC software cannot accommodate separating columns or labelling. A postal pen-and-paper- CBC will be designed

4.5.3 Discussion of Developmental study 2.

Overall, a 'think aloud' based discussion group was useful in developing respondent efficiency. Many other relevant attributes emerged and discussion around the layout and comprehension of the ACBC task emerged. The results reinforce some of the findings of the previous developmental study as well as illuminating some new issues. For example, the importance of clarifying the instructions and reiterating that this is a hypothetical task emerged again.

The selection of attributes and levels

In the 'priority worksheet exercise' respondents tended to include the attribute and levels that had been included in the ACBC exercise. This was encouraging, though it may also be evidence of respondents using an availability heuristic, i.e. that respondents might have relied upon attributes that are available (the ones already presented to them) rather than consider other alternatives. Respondents suggested a vast range of additional attributes from treatments beyond simple painkillers (like referral for an X-ray or to pain clinic) to the distance to your GP and approachability of the GP. It was decided that using GP treatments and services need to be considered based on NICE guidelines and expert opinion. This will be addressed for the next developmental study.

Respondents felt that the included attributes needed clarifying and rewording. Particularly important was incorporating an element of interference of pain with daily life, mobility, ability to work and perform certain tasks needs to be included within the pain attribute in order to ensure clarity. The 'likelihood of pain relief' attribute (Caldow et al. 2007) was unclear and needs clarifying in terms of whether it is what patients want, expect or what the doctor can offer. Also, within this attribute, the levels were percentages, for example 30% pain relief, and these

were considered clinical and meaningless and it was suggested to include percentages in brackets after description. One respondent queried whether this 'likelihood of pain relief' attribute can be combined with 'pain compared to usual' in terms of what is driving the person to go and what can be expected on arrival. One respondent identified a dichotomy between friendly and businesslike and informal and businesslike. It is important that the levels are realistic and plausible and a more clearly defined attribute that may replace this one is the GP's attitude, an attribute also identified by respondents as salient.

The significance of the influence of illness perceptions (i.e. cause and curability) of osteoarthritis and perceptions of health professionals upon the decision to consult was highlighted. These will be included in the subgroup analysis criteria as they are clearly influencing factors and may help to get a deeper insight into the choices that are made.

The construction of the conjoint tasks

A key finding of developmental study 2 is that a full-profile approach (all six attributes presented at same time) is not appropriate for this context (i.e. determinants of GP consultation). There is too much complex information on the screen at one time for respondents to hold in their memory and process effectively (see Table 4.1, on page 103). This finding is consistent with the work in cognitive psychology which suggests that the capacity for processing information in individuals is limited to seven (+/- 2) pieces of information at one time (Miller 1956). Respondents were also overwhelmed by the demands of the task and the amount of information (as well as the ACBC layout not being labelled) compounded this. This clear need for a limited amount of attributes in each conjoint task is a critical finding and suggests a partial-profile design (presenting

four attributes at one time) is more appropriate (than a computer-based ACBC format). Respondents are clearly able to make trade-offs and it is unclear whether it is necessary to include opt-out profiles, as this may allow them to refrain from making trade-offs and avoiding thinking deeply about their actual priorities.

One respondent felt that if the individual profiles were on separate pages it would be clearer to understand. In order to avoid overburdening respondents and thus threatening validity it may be more appropriate to have only 2 (pairwise) profiles.

Respondents demonstrated that they were willing to make trade-offs (against pain for example), based on an underlying rationality (or perception that the health professional cannot do anything to help). It may be appropriate to include a feedback section asking how respondents completed the questionnaire in order to see how much variation there is in approaches to this survey.

The data collection plan

Overall, respondents were confused and frustrated by this ACBC task and it was clear that this was because the format was over burdensome and the exercise was not sufficiently explained. The central question ‘what will encourage you to consult the GP’ is ambiguous and needs rewording in simple language and repeated consistently at every page. The ‘think aloud’ volunteer consistently failed to grasp the hypothetical nature of the task. Instead, he superimposed his personal experience of previous consultations and his expectations onto the tasks. Moreover, the similarity of the choices tasks made the task become repetitive for respondents. Respondents thought that in order to capture the most accurate answer the task must be reiterated at all possible opportunities, including the answer section.

A key finding is that all respondents disapproved of a web based/computer task. Respondents felt that because computer literacy and accessibility of computers for people in the target age range may be limited a computer-based exercise would be an inappropriate data collection procedure. One respondent suggested that the computer option limited their ability to go back and change answers and that a traditional pen-and-paper method to be completed in the comfort of their own homes. Rather than coming into the university to complete an ACBC, a paper-based questionnaire that can easily be filled out at home by respondents would be more appropriate (and familiar) for people in the target age range.

4.5.4 Implications for main study design and further developmental work

The decision to consult the GP for joint pain is complex and this test revealed how many attributes can characterise this decision, as well as crucial information about how much information can be processed effectively at one time. The feasibility of computerised methods of CA for the over 50 population is unclear. The following issues in the development of the questionnaire are highlighted.

- Need for a 'demonstration of how to fill out the choice tasks' at the beginning.
- Need to repeat 'this is a hypothetical task' throughout.
- A design presenting only four attributes (partial-profile- something unique to Sawtooth Software) and 2 profiles (scenarios) at a time (and incorporating labelling where possible) may be more appropriate.
- A postal pen-and-paper-based CBC may be more appropriate than ACBC.

- Consider using NICE guidelines to identify all treatments: From painkillers only to a promising 'new' treatment and referral for an X-ray. Seek an expert medical opinion where possible. Change to treatments and services (based on NICE guidelines) rather than pain relief.
- Include 'disruption of pain to your everyday life compared to usual' as an attribute.
- Consider 'other health problems' and ' GP's attitude/manner' as attributes.
- The number of holdouts traditionally used is two, the importance of these needs to be considered further as they significantly add to cognitive burden of respondents.

4.6 Developmental study 3. finalising the design

4.6.1 Method

Design

Developmental study 3 involved testing a choice-based conjoint (CBC) based on attributes and levels prioritised in the findings of the previous tests. The findings from developmental studies 1 and 2 suggest that choice-based methods with two alternatives/profiles (pairwise comparison) per choice task are more acceptable to for people in the target age range and the aim was to confirm that the format was sufficiently simplified. This phases also aimed to confirm whether or not the chosen attributes and levels were all relevant and to confirm whether or not paper-based or computer mode of administration should be used. It was also important to investigate how understandable the self-devised demonstration task was (see Appendix 5a on page 354), whether it was acceptable to remove the IF from each of the choice tasks and how acceptable the feedback section was. This refers to a section in the questionnaire which asks for feedback on the difficulty of

completing the questionnaire and whether there was any other relevant attributes that were missing (see Appendices 5a and 5b).Having found that the computerised version of ACBC was not acceptable it was now important to establish that the pen-and-paper method was preferable.

The CBC design was based on eight attributes with a total of twenty-three levels (3*3*3*5*2*2*2*3=3240 profiles). These were derived from the priorities of the respondents in developmental study 2 and the literature identified. The combination of potential profiles was again reduced to reduce burden for respondents. Four attributes were clinical need factors (pain type, pain severity, comorbidity and impact on everyday activities), one focused on GP manner, one on GP attitude (see Table 4.8 overleaf) and two focused on primary care services (GP treatments and GP follow-up).

The questionnaire consisted of sections A to D. Section A included an introduction and general instructions, section B was a rating task (from 1-10, with 10 being most important) that included eight attributes that may influence decisions to consult the GP for joint pain. Section C involved four parts: firstly, some prescriptive instructions on ‘how to complete the questionnaire’, secondly a simple paragraph to outline how to fill in the questionnaire, thirdly the choice tasks themselves and fourthly, a feedback section on the choice tasks. Section D included respondent demographics, a coping and illness perceptions questionnaire (see Appendix 5a). The aim was to compare the first two parts of Section C to see which respondents preferred: a simple paragraph or prescriptive instructions about how to fill in the questionnaire. There were sixteen choice sets, and each involved respondents choosing between two scenarios (made up of four attributes) as to which would most likely lead them to consult a GP for joint pain.

Table 4.8: Attributes and levels used in developmental study 3.

Attributes		Levels		Basis	Andersen-Newman (1995) model categorisation
1	Pain type	1.1.	I am experiencing a dull aching pain, which is constant over time	(Hopman-Rock et al. 1997, Sanders et al. 2002, Hawker et al. 2008)	Need factor
		1.2.	I am experiencing short episodes of a more intense, often unpredictable pain		
2	Pain Severity	2.1.	The intensity of this pain is mild	(Ratcliffe et al. 2004, Hawker et al. 2008, Hagen et al. 2000)	Need factor
		2.2.	The intensity of this pain is moderate		
		2.3.	The intensity of this pain is severe		
3	Impact on everyday activities	3.1.	The pain is not disrupting my everyday living	(Thorstensson et al. 2009, Sanders et al. 2002, Mitchell et al. 2006, Rosemann et al. 2006)	Need factor
		3.2.	The pain is disrupting some everyday living		
		3.3	The pain is disrupting most everyday living		
4	GP's manner	4.1.	The GP will seem to be rushed	(Morgan et al. 2000, Rosemann et al. 2006)	Enabling factor
		4.2.	The GP will take time to listen		

Table 4.8 Cont: Attributes and levels used in developmental study 3.

Attributes		Levels		Basis	Andersen-Newman (1995) model categorisation
5	GP attitude	5.1.	The GP will regard my joint pain as part of the normal ageing process	(Sanders et al. 2002, Gignac et al. 2006, Jinks et al. 2007)	Enabling factor
		5.2.	The GP will regard my joint pain as more than just part of the normal ageing process		
6	GP follow-up services	6.1.	The GP will offer me a referral to a hospital specialist (rheumatologist, orthopaedic doctor, pain clinic)	(Tallon et al. 2000, Sanders et al. 2002, Mann & Gooberman-Hill 2011)	Enabling factor
		6.2.	The GP will offer me a referral to a physiotherapist		
		6.3.	The GP will offer me a referral for an X ray		
7	GP treatments	7.1	The GP will offer me Painkillers only	(Tallon et al. 2000, Sanders et al. 2002, Mann & Gooberman-Hill 2011)	Enabling factor
		7.2.	The GP will offer me an Intra-articular injection		
		7.3.	The GP will offer me a promising 'new' treatment		
		7.4.	The GP will offer me lifestyle advice		
		7.5.	The GP will offer me alternative complementary or alternative medicines and treatments (acupuncture)		
8	Other health problems	8.1	I have no other (minor) health problems	(Bedson et al. 2007, Kadam & Croft 2007)	Need factor
		8.2.	I have other (minor) health problems		
		8.3	I have other (major) health problems at the moment		

Participants

As for developmental study 2, members of the ARUKPCC's RUG were invited to participate in this study by postal invitation. There were 5 females and 6 males (all aged 50 years or over).

Data collection procedure

After respondents had agreed to participate, one cognitive interview (within a wider discussion group) was conducted on the university premises. Participants were informed in a postal information sheet about the nature and type of questions that they might expect to be asked (see Appendix 5c).

One respondent was asked to complete the tasks (see Appendix 5a) in the presence of the researcher and members of the discussion group. One respondent was asked to volunteer to 'think aloud' for the group. One 'think aloud' CI was conducted based on a paper-based CBC within a group of 11 members of the RUG. After each page of the CBC task was completed, a five to ten minute group discussion was instigated by researchers (DC, CJ, MF). Issues that arose during the previous page were discussed. The sessions were tape-recorded and transcribed verbatim. Respondents were offered an opportunity for expenses to be paid in full.

Analysis

As with developmental study 1 and 2, the data was transcribed and examined to identify reoccurring and dominant themes across interviews. Data was then analysed according to the aims of the developmental studies set out in section 4.3 on page 106.

4.6.2 Results of developmental study 3

See Appendix 5d for the transcript from developmental study 3. Table 4.9 on the following page presents the results from the group discussion.

Table 4.9: Results of developmental study 3

Objectives (1-3)	Important findings	Transcript quotations	Potential revisions
1. Attributes and levels	'Other health problems' considered relevant	*'Yes I do actually because other problems can be contributory' (Respondent male 1)* *'I think for most people it will be a psychological factor' (Respondent male 1)* *'Because people are human and we tend to sort of waffle off the main subject and think ooh while I'm here I'll just mention so I think it has relevance, it was in my thinking as I went through any questionnaire' (Respondent female 1)*	'Other health problems' to be included as an attribute
1. Attributes and levels	GP attitude and GP manner may need merging	*'Looking at G and H, the GP's manner is G and the attitude to joint problems is H, I think those two are likely to be fused together' (Respondent female 1)* *'Approachability I know is an old fashioned word but it's one I think older people will relate to' (Respondent female 1)*	Merge GP attitude and GP manner together, or pick one) to avoid overlap
1. Attributes and levels	Treatments were acceptable and relevant	*'I thought you have selected those very well and I thought those were about the right number' (Respondent female 1)* *'I had a lot of painkillers and the new treatment idea would appeal so I would definitely include that in your final factors' (Respondent female 1)* *' I must say that my pain has been going on for so long and is understood that I tended to disregard the pain throughout the whole of the filling in of this questionnaire' (Respondent female 1)*	Use 'new treatment' level
2. Construction of conjoint tasks	16 choice sets was too demanding	*'The frames were too repetitive for my concentration (which is very poor). I would suggest 10 frames' (Respondent female 1)*	Reduce to optimum 10 choice sets
2. Construction of conjoint tasks	Reduce words in choice sets	*'I would compress your vocabulary as well as to reduce your number of frames' (Respondent female 1)* *'I think yes I'd do it on the first frame and then put a note at the bottom, IF applies to the other frames' (Respondent female 1)*	Include 'IF' in first choice set and then cut it out for the remaining choice sets

Table 4.9 cont: Results of developmental study 3

Objectives (1-3)	'Think aloud' response errors	Transcript quotations	Potential revisions
2. Construction of conjoint tasks	Respondents change preferences throughout tasks	*'My choice also changed as I went through, my way of thinking and selecting that choice changed as I went through because I thought in this questionnaire, she's playing games with me here, so you couldn't really be. One is I was choosing physio out of that and then I was rejecting physio because something else had cropped up so I thought there is something subtle at work here' (Respondent female 1)*	Do not include holdouts to test for rationality as it is clear that people change their choices throughout the questionnaire
2. Construction of conjoint tasks	Respondents felt limited to choose only one approach to decision-making	*'I can't tick one of those boxes because in some I did one and some I did another and some I did another again. I have used at least about four of those throughout the system, so if I tick one box, what will be the value of that to you? It could lead you to feel like you've got a secure result when you haven't' (Respondent male 3)*	In feedback section allow respondents to tick more than one box to cater for range of techniques used
3. Data collection plan	Clarity of instructions needed	*' well instead is likely to 'expect to offer' or 'think might be offered' something like that' (Respondent male 2)* *'will be able to offer you' (Respondent female 2)*	Make it clear that the GP <u>will</u> offer you investigations/treatments
3. Data collection plan	A prescriptive demonstration was complex	*'The paragraph is better yes, its shorter, easier' (Respondent male 1)* *' It's a better idea …leaving it free for people to use their own minds and which way they read that.(Respondent female 1),* *'I think that last line is important 'it is entirely up to you' how you do it' (Respondent female 2)*	Remove warm up demonstration and explain simply that people use different methods
3. Data collection plan	Clarity of instructions needed: e.g.	*'I'd put, instead of putting a number, you should put please circle one number' (Respondent female 1)*	Change to 'please circle one number' to ensure understanding
3. Data collection plan	The term 'hypothetical' was too technical and unfamiliar to respondents	*' I don't want to sound condescending but some of the public will not know what hypothetical means' (Respondent female 1)*	Use 'imaginary' as it is easier to understand

4.6.3 Discussion of developmental study 3

Overall, the 'think aloud' based discussion group technique proved useful in finalising some of the aspects of developing respondent efficiency. However, there were a vast range of potentially salient attributes that arose from all three of the developmental studies and it was important at this final stage to weigh up the suggestions of respondents, (i.e. which attributes were most acceptable and which needed merging or rewording). The layout and comprehensability of the task was vastly improved at this stage. The results reinforce some of the findings of the previous developmental study as well as illuminating some new issues. For example, the importance of clarifying the instructions and reiterating that this is a hypothetical task was reinforced, as well as a pen-and-paper-based CBC being preferable to the RUG. Some suggested amendments to different aspects of the questionnaire were addressed as a result of this developmental study. These are discussed below in accordance with the results in Table 4.9.

The selection of attributes and levels

Respondents confirmed that other 'health problems' is a relevant and plausible attribute to be included, indicating that the minor and major categories were easily understood. It was suggested that the GP attitude and GP manner overlapped, and so one had to be chosen. Approachability of GP was suggested by one respondent, but it was decided that the GP attitude encapsulated something deeper in terms of the treatment offered by the GP. From the systematic review it was clear that it was not an attribute that had been used before, so it was considered to be an innovation of a patient-centred care attribute (and one pertinent to the decision to consult the GP as identified in the literature review in chapter 1).

Respondents confirmed that the GP treatments and services - based on NICE guidelines– were relevant and plausible, and the ‘promising new treatment’ level was considered desirable. With reference to the levels included within GP treatments and investigations, though these all were agreed to be relevant, it was decided at this point to consult and gain knowledge from a research GP who has expertise in requirements of a consultation (Louviere et al. 2000). The research GP suggested that the attribute levels for both GP treatments and investigations should reflect: 1) the minimum content of a GP consultation for joint pain/OA – i.e. GP determined 'routine' care (or just 'routine care'), 2) an optimal guideline care consultation, i.e. guideline directed 'best practice' care and 3) consultation that may exceed even optimal guideline care, i.e. beyond 'best practice' care. It was suggested that these 3 phrases portray a sense of increasing levels of input and care structure, whilst suggesting to the reader a general feeling of what that care level entails.

The pain characteristics and pain disruption attributes had been prominent in the literature and prioritised by the RUG throughout all of the developmental studies. It was therefore decided to include these, alongside comorbidity. This meant that three clinical need attributes (pain disruption, pain characteristics and comorbidities) and three service attributes (GP attitude, GP assessments and primary care management options) were selected. This meant that there was a balanced design and might enable a conclusion to be drawn about which factors overall are the most important to respondents.

The construction of the conjoint tasks

Developmental study 3 highlighted that sixteen choice sets (even with only four attributes presented at one time) was too cognitively demanding for people in

the target age range. It was decided to reduce the questionnaire to a maximum of 10 choice sets (less than the recommended number of 12, by Ryan et al. 2001) with three attributes per scenario.

One respondent felt that if the individual profiles were on separate pages it would be clearer to understand. In order to avoid overburdening respondents and thus threatening validity, it was decided that only 2 (pairwise) profiles would be used in the final design.

Data collection plan

Overall, respondents were happier and more accepting of the paper-based partial-profile CBC design and it was clear that this was because it was better explained, simplified and in an easier to digest format. The 'demonstration of how to fill out the choice tasks' was considered to be confusing and 'directing thought' and was suggested to be removed. Instead RUG members advised to include a paragraph on how people may complete a CBC, leaving it free for people to use their own minds and emphasising 'it is entirely up to you how you answer it'.

It was suggested that the instructions could be improved in the following ways: change the word hypothetical to 'imaginary', include 'IF' in first choice set and then cut it out for the remaining choice sets, insert 'please circle one number' in the rating exercise to ensure respondents understand the instructions and make it clear that the GP will offer you investigations/treatments. In the feedback section it was suggested that respondents should be allowed to tick more than one box to cater for a range of techniques used. Moreover, this has implications for holdouts, suggesting that they are not necessarily needed to test for rationality as it is clear that people change their choices throughout the questionnaire.

A paper-based partial-profile design (presenting three attributes at one time) was preferred as it was considered more appropriate for people in the target age range. This design will be used in the main study (and since this format is unique to Sawtooth Software, this is the software that will be used).

4.7 Summary of findings

Test 3 confirmed that the salient attribute levels, format and length of questionnaire, and completion instructions needed final amendments in order to be acceptable to respondents before fielding. The pen-and-paper format was clearly preferred over the computerised version. The following issues in the development of the questionnaire for the main study design are highlighted.

- The choice of attributes and levels developed so far were all considered relevant and plausible. The significance of 'other health problems' to be included as an attribute was confirmed and will be included.
- The GP attitude and GP manner were considered to be overlapping so one attribute (the GP attitude) was selected as the most salient.
- The number of holdout choice tasks traditionally used is two and the inclusion of these, when considered in terms of the cognitive burden to respondents, was considered unnecessary. Holdout choice tasks will not be included in the final design.
- The instructions will be made clear and repeated where possible.
- The IF' will appear in first choice set and then cut out for the remaining choice sets.
- After a consultation with a research GP who has expertise in requirements of a consultation (Louviere et al. 2000) it was decided that in the final design the attribute levels for both GP treatments and investigations may

reflect: 1, the minimum and determined 'routine' care, 2, an optimal guideline care consultation and 3, beyond 'best practice' care.

- The full list of all attributes identified and suggested (in Appendix 6) were collated and considered in light of the findings of developmental studies. These were reduced down to six attributes covering sixteen levels in total.
- Three clinical need attributes (covering 8 levels) were chosen: including pain disruption, pain characteristics and comorbidities. Three service attributes (covering 8 levels) were selected: including GP attitude, GP assessments and primary care management options. Incorporating three service (enabling) and three clinical need factors meant that there was a balanced design and might enable a conclusion to be drawn about which factors overall are the most important to respondents (for example, the service factors had a relative importance of 50%).

Table 4.10: Final list of attributes and levels to be included in final CBC questionnaire

Attribute level				Andersen-Newman (1995) model categorisation
1	Pain characteristic	1.1	Dull aching pain	Need factor
		1.2	Severe unpredictable episodes	
2	Level of disruption to everyday life	2.1	None	Need factor
		2.2	Some	
		2.3	Most	
3	Competing comorbidity	3.1	None	Need factor
		3.2	Minor	
		3.3	Major	
4	GP assessment/ investigations	4.1	Subjective only	Enabling factor
		4.2	Subjective-thorough physical exam	
		4.3	Subjective-thorough physical exam-X rays/blood tests	
5	Primary care management/treatment	5.1	Verbal advice-prescribed analgesia	Enabling factor
		5.2	Written advice-prescribed analgesia-PN follow-up-physio referral	
		5.3	Written advice-prescribed analgesia-PN follow-up-physio referral-promising new treatment	
6	GP attitude	6.1	Normal ageing process-accept it	Enabling factor
		6.2	Legitimate health problem-requires treatment	
GP General practitioner, PN Practice Nurse				

4.8 Discussion

Overall, the combination of methods (literature reviews, research expert opinion, 'think aloud' cognitive interviews and discussion groups) proved useful methods for developing respondent efficiency of the conjoint questionnaire. These methods ensured that the selection of salient attributes and levels, construction of conjoint tasks and data collection plan were appropriate for the target population. However, there were some limitations of using cognitive interviewing with choice-based conjoint analysis. The open-ended nature of the 'think aloud' protocol caused two problems, reflected in previous research. Firstly, there was a tendency for subjects to lose focus and drift from the exercise in hand. It was important to refocus them, for example, 'remember to tell me how you got to that decision' (Campanelli 1997). Respondents tended to want to please interviewers using phrases like 'it's perfect' (female, 75), despite clear evidence that respondents struggled to make trade-offs. Similarly, two respondents seemed to mask their true feelings, for example, 'we won't be nasty' (female, 80). This raises the issues of social desirability and researcher presence as potential threats to validity.

A significant challenge was that developmental study 1 was not entirely successful because one participant effectively refused to take part. Due to this negative experience of one person refusing to participate in developmental study 1 the focus for the remaining developmental studies was to reduce and simplify the choice tasks and information presented to respondents as much as possible. It appeared that CIs are more effective when done as a one-on-one method rather than group-based because there is less influence from others upon respondents' individual answers.

Using a ‘think aloud’ protocol provided evidence that supports previous research into the behavioural limitations of CBC. These include the influence of personal experience and the capability and inclination for patients to make trade-offs when completing a CA questionnaire (Cheraghi-Sohi et al. 2007). The open-ended ‘think aloud’ format proved beneficial, allowing one respondent to shed some light on an under researched area, the influence of emotional processing upon choice behaviour. CIs when used in the formative development stage of a CBC can inform the development of more effective attributes and levels. This technique helped particularly with the comprehension of the wording and to define the attributes and concepts clearly.

Over the course of the developmental studies a large number of potentially relevant, often complex attributes were identified and suggested by respondents (in Appendix 6). In the end a final decision was made to reduce the extensive list of attributes down to six (covering sixteen levels). The six attributes selected for the final design had all been acceptable to the RUG (although not in exactly the same format, for example, the 3 phrases of care structure in the GP assessment and primary care management attributes appeared differently). This was decided in order to ensure a balanced design of three clinical need and three service related attributes.

However, the extensive list of possible determinants (see Appendix 6) raises further issues. It may be that such a large list could be accommodated with adaptive computer CA designs though it was clear that the RUG respondents were not in favour of using a computer-based method, nor where they able to manage so much information at one time. This implied the need to prioritise and select attributes and levels for pen-and-paper administration. However, even if these

were prioritised to people's general satisfaction the assumption of ceteris paribus (that respondents will hold equal all other attributes not included in the partial-profile choice sets) may be difficult given that each individual may have other determinants in their mind when completing the choice task. It was decided to prioritise and select attributes and levels for pen-and-paper administration for the final design and the assumption of ceteris paribus was something that had to be accepted as a necessary part of selecting a partial-profile design – a design chosen ultimately to avoid overburdening respondents with too many attributes to consider.

Overall an iterative approach was used with data collection. This was based on recommended methods in the field, including a literature review, cognitive interviews, research expert opinion and extensive developmental testing with the RUG. There was a tension between the aim in qualitative work to explore and describe, and the selectiveness needed to encapsulate a minimum number of attribute levels salient to the decision to consult the GP for joint pain.

It is important to acknowledge that these developmental studies were rooted in the field of patient and public involvement (PPI) in research, something that in the last decade has become a requirement in applied health research and compulsory for funding bodies (Ives et al. 2012). The findings in this chapter are consistent with research into PPI which suggests that PPI has a clear impact upon the research process in the early stages of research, assisting in research design such as developing research questions, identifying and selecting topics and improving the comprehensability of research language in questionnaires (Brett et al. 2010).

The final list of attributes is presented in Table 4.10. The final design of the CBC (including all accompanying literature) is presented in chapter 6 (section 6.3.5). Chapter 5 will now look at the developmental studies that were undertaken in order to ensure statistical efficiency, before the main study design is presented in chapter 6.

5 Chapter Five: Methods II: developmental studies: statistical efficiency

5.1 Introduction

In designing a CBC study to investigate the determinants of GP consultation for joint pain, Chapter 4 considered the perspective of a Research Users' group in the context of existing literature and optimal design of CA studies. This work resulted in setting certain design features, namely: the use of 6 attributes (two 2-level attributes, four 3-level attributes), the use of partial (based on 3 attributes) rather than full-profiles (due to the number and complexity of the attributes), the presentation of only two profiles per choice task, and 10 choice tasks feasible for completion. Postally-administered pen-and-paper questionnaires were chosen (due to unpopularity and logistical issues of computer-based methods). The total number of full-profiles that could be obtained from this number of attributes and levels is 324 (2*3*3*3*3*2).

In this chapter the focus is on an evaluation of the statistical efficiency of the experimental design (Bridges et al. 2011), using the Advanced Design Module (ADM) within Sawtooth Software (Inc. Orem, UT), and the use of sensitivity analyses to finalise the experimental design.

5.1.1 Experimental design

The experimental design is defined as *'the process of systematically manipulating the attribute levels to create the profiles and tasks'* (Bridges et al. 2011). Within Sawtooth Software (v7.0) there are a number of design parameters required to generate a pen-and-paper-based (PPCBC). The software minimises the confounding of attribute main effects by selecting a subsample of hypothetical product combinations based on the statistical principles of *'orthogonality (no*

correlation between attributes), level balance (each attribute level occurs with equal frequency) and minimal level overlap (each attribute level only appears once in a given choice)' (Bridges et al. 2009, p.11).

This study selected a design with pairwise (two scenarios/situations) per choice task with each situation described by three attribute levels (see Table 4.10 on page 151). The aim was to create a series of tasks that would generate as much statistical information as possible to obtain unbiased, precise parameter estimates of the preference model (Bridges et al. 2011).

The design of choice tasks involved finding a balance of statistical efficiency and respondent efficiency. Statistical efficiency relates to how precisely a given set of conjoint questions can estimate the parameters of interest and respondent efficiency relates to ensuring that an unacceptable cognitive burden is not placed on respondents when filling out choice tasks. Statistical efficiency is increased when the number of choice tasks, number of attributes per scenario or the number of questionnaire versions is increased (Orme 2006). However, increasing these aspects reduces respondent efficiency (Cunningham et al. 2009). This finding was replicated in the preliminary developmental work in chapter 4. It was also decided that attributes appear in random order within the profile, to minimise order effects.

The number of questionnaire versions posed a more practical challenge as the study was intended to be administered in pen-and-paper format. The design, printing, and keeping track during mailing of dozens of different questionnaire versions was to be avoided if possible. Increasing the number of versions offers more combinatorial variation though the advantages of increased variation are minimal (Orme 2010).

A further consideration was whether to limit the focus to main effects or to attempt to estimate interactions. These are defined as an occasion when *'the levels from two attributes combine to create something better or worse than their independent values might suggest'* (Orme 2006*,* p.132). In this study, interactions seemed plausible. For example, a pain characteristic level of 'severe unpredictable pain' and pain disruption level 'pain is disrupting some of my everyday life' might combine to create something worse than their individual values. However, partial-profile designs are underpowered and inefficient at estimating interactions because they may only occur together within a profile on a very small number of occasions.

The Advanced Design Module (ADM) within Sawtooth Software (version 7.0) provides an opportunity for sensitivity analyses using simulated data.

5.2 Evaluating statistical efficiency using simulations in Advanced Design Module

The Advanced Design Module (ADM) estimates the precision of the parameter estimates based on the combined elements of design efficiency and sample size (respondents x tasks) using an aggregate Multinomial Logit (MNL). MNL is a general model for fitting choice data, but here refers to a specific pooled (aggregate) logit model (see Appendix 9 for more information). To perform the test, the researcher must supply information on the expected number of respondents. Based on this information the ADM simulates random (dummy) respondent answers and generates two, related measures of the statistical efficiency (precision) for the model overall - a 'strength of design' estimate and the D efficiency statistic (Orme 2010) – and estimates of the standard errors for individual attribute levels. An illustration of the typical output from an ADM

simulation is provided in Figure 5.1 based on a simulation of a partial-profile CBC design with 12 choice tasks, 500 respondents, and estimating only main effects (i.e. no interactions).

Figure 5.1: ADM output for testing statistical efficiency

Attribute level				Effect	SE
1	Pain characteristic	1.1	Dull aching pain	0.00807	0.01827
		1.2	Severe unpredictable episodes	-0.00807	0.01827
2	Level of disruption to everyday life	2.1	None	0.01591	0.2987
		2.2	Some	0.01974	0.2983
		2.3	Most	-0.03565	0.2985
3	Competing comorbidity	3.1	None	0.3866	0.02985
		3.2	Minor	-0.05298	0.02986
		3.3	Major	0.01432	0.02986
4	GP assessment/ investigations	4.1	Subjective only	-0.01790	0.02984
		4.2	Subjective-thorough physical exam	-0.01432	0.02985
		4.3	Subjective-thorough physical exam-X rays/blood tests	0.03223	0.02986
5	Primary care management/treatment	5.1	Verbal advice-prescribed analgesia	0.03177	0.2987
		5.2	Written advice-prescribed analgesia-PN follow-up-physio referral	-0.04027	0.02987
		5.3	Written advice-prescribed analgesia-PN follow-up-physio referral-promising new treatment	0.00849	0.02985
6	GP attitude	6.1	Normal ageing process-accept it	-0.01509	0.01827
		6.2	Legitimate health problem-requires treatment	0.01509	0.01827
GP General Practitioner, PN Practice Nurse Choice tasks (n=12), respondents (n=500), questionnaire versions (n=30), The strength of design =1532.67					

The effects are the estimated partworths for the 6 attributes and the 16 levels. These are insignificant in the ADM as the logit is based on random data. The SE (aggregate standard error) column reflects the precision of the partworth estimates for each attribute. A lower SE represents greater precision (Orme 2010).

It is recommended that standard error for main effects should be no larger than 0.05 (acceptable) and ideally less than 0.025 (Orme 2010). In the example above the SE for all attributes are acceptable. The estimates for two-level attributes are more precise (.018) than three-level attributes (.029), because each three-level appears in the choice tasks fewer times.

The strength of design for this model reflects the attribute standard errors, with higher values representing a more statistically efficient design (although there appears to be no standard interpretation of the absolute values of this figure). The d-efficiency (not shown above) is based on the strength of design statistic and calculated from the ratio of strengths of design for two models (i.e. statistical efficiency of one design relative to another) (Orme 2010). For example, if a simulation based on the same features as above was ran but included only 11 choice sets (and the strength of design for that model was estimated at 1379.40) the d-efficiency value would be 0.90; we might say that including 11 choice sets is 90% as efficient as including 12.

5.3 Aims and objectives

The aim of this chapter was to finalise the design of the PPCBC design by performing sensitivity analyses on the statistical efficiency of the design. This was done by simulating different levels of certain key features in the design, using the ADM to ensure that partworths are estimated with adequate precision prior to finalising the design.

The specific objectives were:

1. To evaluate the effect on statistical efficiency (strength of design, d-efficiency, standard error of main effects per attribute) of changes in the number of choice sets.

2. To evaluate the effect on statistical efficiency of changes in the number of questionnaire versions.
3. To evaluate the effect on statistical efficiency of changes in the number of interaction effects.

5.4 Methods

All simulations were carried out by the author on the ADM, prior to finalising the survey instrument. In each simulation the following design features were held constant: number of respondents = 500; partial-profile; two profiles per choice task; 6 attributes (two 2-level, four 3-level).

5.4.1 Simulation 1: Number of choice sets

A general rule of thumb is that every level appears three to six times to each respondent (Orme 2010). In the current study, for each of the 16 levels to appear a minimum of three times (16*3=48) and given three attributes per scenario alternative and two scenarios per choice task (48/6= 8) it was calculated that 8 choice sets was the minimum amount of choice sets required for statistical efficiency during developmental studies with the RUG. It was established that 10 was their preferred number of choice tasks but it is recommended that no more than 12 choice sets be presented in total (Ryan et al. 2001). Simulations were therefore ran on 8 to 12 choice tasks, keeping all other design features constant.

5.4.2 Simulation 2: Number of questionnaire versions

Advice from Sawtooth Software recommends that the number of random choice tasks multiplied by the number of questionnaire versions is greater than or equal to 80. Assuming 8 choice tasks (minimum from above) this would suggest a

minimum of 10 questionnaire versions. The simulation evaluated statistical efficiency of the design when using between 10 and 100 questionnaire versions.

5.4.3 Simulation 3: Number of interactions

To evaluate the potential to estimate precisely a limited number of interactions, the following interactions were included: Pain Characteristics* Comorbidity, Pain Disruption* Comorbidity, Pain Characteristics * Pain Disruption, + Pain Disruption* Comorbidity. These were chosen to represent a 2-level * 3-level attribute interaction, a 3 * 3, and a 2 * 3 plus 3 * 3 interaction respectively.

5.5 Results

5.5.1 Simulation 1: Number of choice sets

Table 5.1: Effect on statistical efficiency of simulated number of choice sets

No. of choice sets	Strength of design	d-efficiency	SE of effect > 0.025
8	1019.11	66%	disruption to daily activities (0.036); comorbidity (0.036); primary care management (0.036); GP assessment (0.036)
9	1145.71	75%	disruption to daily activities (0.034); comorbidity (0.034); primary care management (0.034); GP assessment (0.034)
10	1276.77	83%	disruption to daily activities (0.032); comorbidity (0.032); primary care management (0.032); GP assessment (0.032)
11	1404.73	92%	disruption to daily activities (0.031); comorbidity (0.031); primary care management (0.031); GP assessment (0.031)
12	1532.67	100%	disruption to daily activities (0.029); comorbidity (0.029); primary care management (0.029); GP assessment (0.029)
Respondents (n=500), questionnaire versions (n=30); partial-profile; two profiles per choice task; main effects only			

Table 5.1 shows that although the strength of design of 10 choice sets (SE = 0.032) was less than with 12 choice sets (SE = 0.029) both were within the acceptable guidelines. All of the sixteen attribute levels were acceptable (SE = <0.05), twelve attribute levels were acceptable but not ideal and four attribute levels (the two level attributes) were ideal (SE = <0.025).

Similarly, in table 5.1 where 12 choice sets set the 100% standard and all others are as a percentage of 1532 (e.g. 11 = 1404/1532), it is clear that by reducing 12 choice sets to 10 there is 17% loss of efficiency.

5.5.2 Simulation 2: Number of questionnaire versions

Table 5.2: Effect on statistical efficiency of simulated number of versions

No. of versions	Strength of design	d-efficiency	SE of effect > 0.025
10	1526.95	99.6%	disruption to daily activities (0.029); comorbidity (0.030); primary care management (0.030); GP assessment (0.030)
20	1528.22	99.7%	disruption to daily activities (0.030); comorbidity (0.029); primary care management (0.029); GP assessment (0.030)
30	1532.67	99.9%	disruption to daily activities (0.029); comorbidity (0.029); primary care management (0.029); GP assessment (0.029)
40	1530.92	99.8%	disruption to daily activities (0.029); comorbidity (0.029); primary care management (0.029); GP assessment (0.029)
50	1532.83	100%	disruption to daily activities (0.029); comorbidity (0.029); primary care management (0.029); GP assessment (0.029)
Choice tasks (n=12), respondents (n=500); partial-profile; two profiles per choice task; main effects only			

Table 5.2 shows that although the strength of design of 10 versions (SE = 0.030) is less than with 50 versions (SE = 0.029) both SEs are within the acceptable guidelines. All of the sixteen attribute levels were acceptable (SE = <0.05), twelve attribute levels were acceptable but not ideal and four attribute levels (the two level attributes) were ideal (SE = <0.025).

Table 5.2 shows that the maximum statistical efficiency of designs is reached when using 50 versions. There is effectively no difference in statistical efficiency in the range of 10-50 versions.

5.5.3 Simulation 3: Number of interactions

Table 5.3: Effect on statistical efficiency of simulated number of interactions

No. of interactions	Strength of design	d-efficiency	SE of effect > 0.025
None	1532.67	100%	disruption to daily activities (0.029); comorbidity (0.029); primary care management (0.029); GP assessment (0.029)
Pain Characteristics* Comorbidity	1058.41	69%	disruption to daily activities (0.029); comorbidity (0.030); primary care management (0.029); GP assessment (0.029) Pain characteristics* Comorbidity (0.083)
Pain Disruption* Comorbidity	770.80	50%	disruption to daily activities (0.029); comorbidity (0.029); primary care management (0.029); GP assessment (0.029) Pain Disruption* Comorbidity (0.10)
Pain Characteristics* Comorbidity+ Pain Disruption* Comorbidity	634.82	41%	disruption to daily activities (0.029); comorbidity (0.029); primary care management (0.029); GP assessment (0.029) Pain characteristics* Comorbidity (0.084) Pain Disruption* Comorbidity (0.104)
Choice tasks (n=12), respondents (n=500), questionnaire versions (n=30); partial-profile; two profiles per choice task			

Table 5.3 shows that the strength of design of no interactions (SE = 0.029) is within the acceptable guidelines. However, when including one interaction (Pain characteristics* Comorbidity) the size of the SEs is no longer acceptable (SE = 0.083) i.e. the SEs are no longer <0.05 and hence are not within the acceptable guidelines.

It is clear from table 5.3 that the strength of design gets lower (and thus SE gets higher) when interactions are included. Table 5.3 reinforces guidelines that suggest that in terms of interactions PPCBC designs are inefficient and likely to be underpowered.

5.6 Discussion

5.6.1 Principal findings

Number of interactions had the greatest effect on statistical efficiency. Therefore, it was decided that interactions would not be included. Number of choice sets had a substantial effect on statistical efficiency, although 10 choice sets (limit for the RUG) was accepted as statistically efficient. The loss of efficiency from 12 to 10 choice sets was considered acceptable (17%) and in order to meet the ideal limit for the RUG the optimum solution was to use 10 choice sets. In terms of holdout tasks (see chapter 2 on page 35), it was decided that these would not be included in the final design of the questionnaire in order to include 10 choice sets and increase respondent efficiency. Results suggest that using 10 versions is acceptably efficient (and will ensure data collection is manageable).

5.6.2 Interpreting the principal findings in the context of previous studies

Of the 15 studies identified in the systematic review in chapter 3 (see Table 3.3 on page 69) three reported a calculated d-efficiency statistic (Cheraghi-Sohi et al. 2008, Longo et al. 2006, Gerard et al. 2008) in advance of conducting their study as a way of informing the design. Similarly Marshall et al. (2010) found that details about efficiency were often missing from the reviewed applications of conjoint analysis in health research. None of the 15 reviewed studies retrospectively reported the statistical efficiency of their design (i.e. back-calculated a strength of design statistic for their study based on actual response). Within the Sawtooth Software (Inc. Orem, UT) literature in marketing research retrospective reporting of the statistical efficiency does not seem to be conventional practice either.

Although there have been few studies in health that have examined interactions, of the 15 reviewed studies four of them tried to estimate interaction terms (Vick & Scott 1998, Scott & Vick 1999, Longo et al. 2006, Gerard et al. 2008). These studies found that interaction effects were illuminating to some degree. One study justified not testing for interactions based on previous empirical studies which suggest that interactions between main effects are negligible (Louviere 1988). In the Sawtooth literature it is suggested that main effects are easier to interpret and mean that there are fewer parameters to estimate (Orme 2006).

Although all of the 15 studies reviewed in chapter 3 (see page 69) used a fractional-factorial design (and none used a partial-profile design) it seems that to estimate main effects only is quite acceptable within the field. For example, in the ISPOR checklist (Bridges et al. 2011) guidelines suggest that researchers should test different design approaches to suit their objectives. It notes that estimating main effects and interactions in full-factorial designs may be too overwhelming for respondents to cognitively manage in a practical way. Bridges et al. (2011) suggest that designs which guarantee that all attribute main effects are independently estimable, and only some interactions are estimable, are commonplace (Bridges et al. 2011). Certainly, in the marketing research on the optimal design of choice-based conjoint designs, researchers tend to focus on the main effects using MNL logit model, and appear to not include interactions between attributes (Kessels et al. 2006).

5.6.3 Strengths and weaknesses

Sensitivity analyses were useful because this allowed the researcher to test the design efficiency, where efficiency is expressed relative to an ideal set of

conjoint questions (design) under consideration. Thus it was possible to consider 10 and 12 choice sets, and test the difference between each design in terms of precision of partworth utility estimates (assuming pooled estimation). However, despite its usefulness in generating an efficient design, overall aspects such as sample framing (to ensure a good response rate and sample size) are ultimately most important to ensure the stability of estimates. This reinforces the importance of ensuring the acceptability of the task to respondents. Kievit et al. (2010) recognised the same issue. Despite having an efficient design they acknowledged that if they had a greater response their estimates would have been more stable (Kievit et al. 2010).

The main limitation is that in terms of estimating interactions PPCBC designs are inefficient and likely to be underpowered. It is therefore not possible to predict (with much precision) whether changes in service provision would selectively encourage those with more severe problems to consult or whether this would equally have the effect of attracting consultations for less severe problems. A full-profile design would be needed. However, the strong feedback from patients in developmental work was that the presentation of 6 attribute levels in this study would not be manageable.

5.6.4 Summary and how results will inform the main study design

The ADM was important for ensuring statistical efficiency prior to finalising the design. The following issues in the development of the questionnaire for the main study design are highlighted:

- The main design will include 10 random choice tasks (excluding holdouts) because the standard errors are in line with the 'acceptable' guidelines and close enough to the ideal guideline.

- The main design will include 10 questionnaire versions. This is above the minimum number of versions (8) and it is less important to D efficiency to include more than 10 versions. Using 10 versions will also keep data collection manageable.
- The main design will look at main effects but not interactions. Results suggest that the PPCBC format does not have the power to precisely estimate interaction effects.

6 Chapter Six: Methods III: main study design, methods, and planned analysis

6.1 Introduction

The overall efficiency of the pen-and-paper-based partial-profile choice-based conjoint (PPCBC) design was developed in chapter 4 (in terms of respondent efficiency) and chapter 5 (in terms of statistical efficiency). This chapter presents the aims and objectives of the study (section 6.2) before presenting aspects of the methods including study design, setting, sample frame data collection procedures and data handling (section 6.3). Plans for statistical analysis (section 6.4) and ethical approval are also presented (section 6.5).

6.2 Aims and objectives

The broad purpose of the study is to understand why some patients with joint pain/OA do not consult their GP despite apparent clinical need, and to ascertain whether (changes in) certain aspects of service provision (would be expected to) facilitate consultation.

The primary aim was to quantify the relative importance of selected clinical need and general practice service factors ('attributes') in the decision to consult the GP in older adults with joint pain/OA.

Secondary aims were:

- To compare the direct ratings of the importance of attributes and the attribute importance scores from conjoint analysis by the same individuals.
- To identify different types of preferences within this heterogeneous population.

The specific objectives were:

- To conduct, analyse, and interpret a cross-sectional survey (postally-administered, self-complete) containing partial-profile choice-based conjoint analysis and pairwise comparisons in a population-based sample of adults aged 50 years and over with joint pain/OA.
- To compare the rank order of attribute importance scores from conjoint analysis and the direct ratings of the importance of attributes.
- To use subgroup analysis to identify different types of consultation preferences amongst older adults with joint pain/OA and conduct between-group comparisons on selected demographic and other patient characteristics.

6.3 Method

6.3.1 Study design

The main study consisted of a partial-profile choice-based conjoint analysis study, based on a cross-sectional survey (single self-complete questionnaire, postally-administered).

6.3.2 Setting

The general population (North Staffordshire).

6.3.3 Sample frame

The sample frame chosen for this study were respondents to a recent 6-year follow-up (6YFU) of a population-based cohort of adults aged 50 years and over, conducted across 6 general practices in North Staffordshire (North Staffordshire Osteoarthritis Project – NorStOP1 & NorStOP2) (Thomas et al. 2004).

All potential participants were aged 50 years and over and registered with one of 6 participating general practices in North Staffordshire when initially invited to join the cohorts in 2002-2003. Respondents will have consented to further contact throughout the duration of the cohort study, including at 6 year follow-up (6YFU).

This sample frame was selected for convenience and to maximise response. Recruiting from a sample 'well-disposed to' research - one used to returning questionnaires to the Arthritis Research UK Primary Care Centre – may ensure a good response rate. It was also chosen for availability of information to provide the basis for rich description and subgroup analysis. This information could then be left out of the conjoint analysis study questionnaire, thereby also reducing the respondent burden of the conjoint questionnaire.

Checks of sample frame

The sample frame was checked for rates of missing data and the likely informativeness. Informativeness was measured by having a heterogeneous sample with a wide distribution across a number of characteristics such as age, gender and joints involved. Checks of the sample frame were done using sociodemographic data, information on joint pain, and psychological factors (Hospital Anxiety Depression Scale and the Brief Illness Perceptions Questionnaire) that had been self-reported by potential participants in their 6-year follow-up questionnaires. The Hospital Anxiety Depression Scale (HADS) is a 14-item self-report questionnaire assessing severity of anxiety and depression, seven items relating to anxiety and seven to depression and each scored on a four-point Likert scale (Zigmond & Snaith 1983). The Brief Illness Perception Questionnaire (BIPQ) investigates cognitive perceptions of illness and contains eight items

scored on a 10-point Likert scale and a ninth open-response item (Broadbent et al. 2006).

A sample of 1563 potentially eligible participants were selected (from NorStOP2 6 year follow-up, plus a random sample from NorStOP1 6 year follow-up). The rates of missing data for the demographic and health related items needed for this thesis were all 5% or below. Within each GP practice there was a consistently higher proportion of females than males in each age group. For all three joint sites (hand hip and knee) a higher proportion of patients had mild pain, followed by moderate pain, and least patients had severe pain. In terms of patterns of joint pain involvement the highest proportion of patients had pain at three sites; hand, hip, and knee (26%) and the lowest proportion of patients had hand and hip pain (7%) (see appendix 7 on page 395). It was decided that the NorStOP cohorts were an informative sample frame for this thesis because they proved to be a heterogeneous sample with a wide distribution across a number of characteristics such as age, gender and joints involved.

6.3.4 Sampling procedure

Individuals were sampled from NorStOP 6-year follow-up respondents according to the following eligibility criteria.

Eligibility criteria:

Inclusion criteria

- Age 50+ years
- Male or female
- Responded to NorStOP1 or 2, 6 year follow-up Health Survey questionnaire
- Reported hand, hip, or knee pain in the past year at 6YFU

- Provided written informed consent to further contact at 6 year follow-up

Exclusion criteria

- Participants who have already been, or were, approached for other studies.
- Deaths and departures since completion of 6YFU Health Survey questionnaire.
- Participants who withdrew from the study after mailout of the NorStOP 6YFU Health Survey questionnaire.
- Participants who appear to be vulnerable groups (e.g. new-onset dementia or severe/terminal illness) after GP screening.

Eligible respondents were sampled (and reduced down to 1170) from NorStOP1 and 2 6YFU for inclusion in the study based on their availability (i.e. they weren't being used in other studies).

6.3.5 Data collection procedures

All potentially eligible participants were identified from anonymised responses to NorStOP1 and NorStOP2 6YFU. Recruitment procedures are summarised in the flow diagram (see Figure 7.1 on page 193). Each step is described below.

Notification of general practices

A Practice Pack, standard for all ARUKPCC studies involving partnership with local general practices, was sent to the participating practices from the Principal Investigator. These contained the following:

- List of project team
- A copy of letter of approval from REC, R&D and practice
- Project Summary

- A copy of study documentation (Questionnaire, Questionnaire Cover Letter, Patient Information Sheet)
- Statement on confidentiality of patient data
- Study protocol

Preparation of mailing database

Prior to mailing, the mailing list of potential participants was checked to ensure that addresses were up to date and to exclude those who had died in the period since their last follow-up.

The list was then screened by the lead GP at the practice to exclude vulnerable groups e.g. new-onset dementia or severe/terminal illness.

Survey mailing procedures

A standard three-stage mailing procedure was used. This procedure has been established for all ARUKPCC surveys.

- Participants were sent a questionnaire *(*Appendix 8a) from the Principal Investigator with a Questionnaire Cover Letter (Appendix 8b) inviting them to take part in the study. This was accompanied by a Patient Information Sheet *(*Appendix 8c*)* outlining the purpose of the study and what participants should do if they want to take part. Participants were asked to complete and return their questionnaire. All participants were given the contact number of a researcher working on the project who would give any other information about the project if needed.
- Non-respondents at 2 weeks - Reminder Postcard (Appendix 8d).
- Non-respondents at 4 weeks – Repeat Questionnaire, Patient Information Sheet, and Repeat Questionnaire Cover Letter (Appendix 8e).

Non-respondents after 6 weeks were assumed to have declined participation and were not contacted again for this study.

Survey questionnaire

Data collection was by single self-complete questionnaire only.

Content: 26 pages*: s*ections on direct ratings, choice tasks, and descriptive characteristics.

Estimated time for completion: 60 minutes.

The content of the questionnaire was selected and specified based on three sources: (i) a review of the published literature (determinants of consultation for joint pain/OA, previous conjoint analysis studies in primary care consultation), (ii) cognitive interviews and focused discussion groups with members of the Keele University Arthritis Research UK Primary Care Centre's Research User Group (RUG), and (iii) discussion within the research team and some expert opinion.

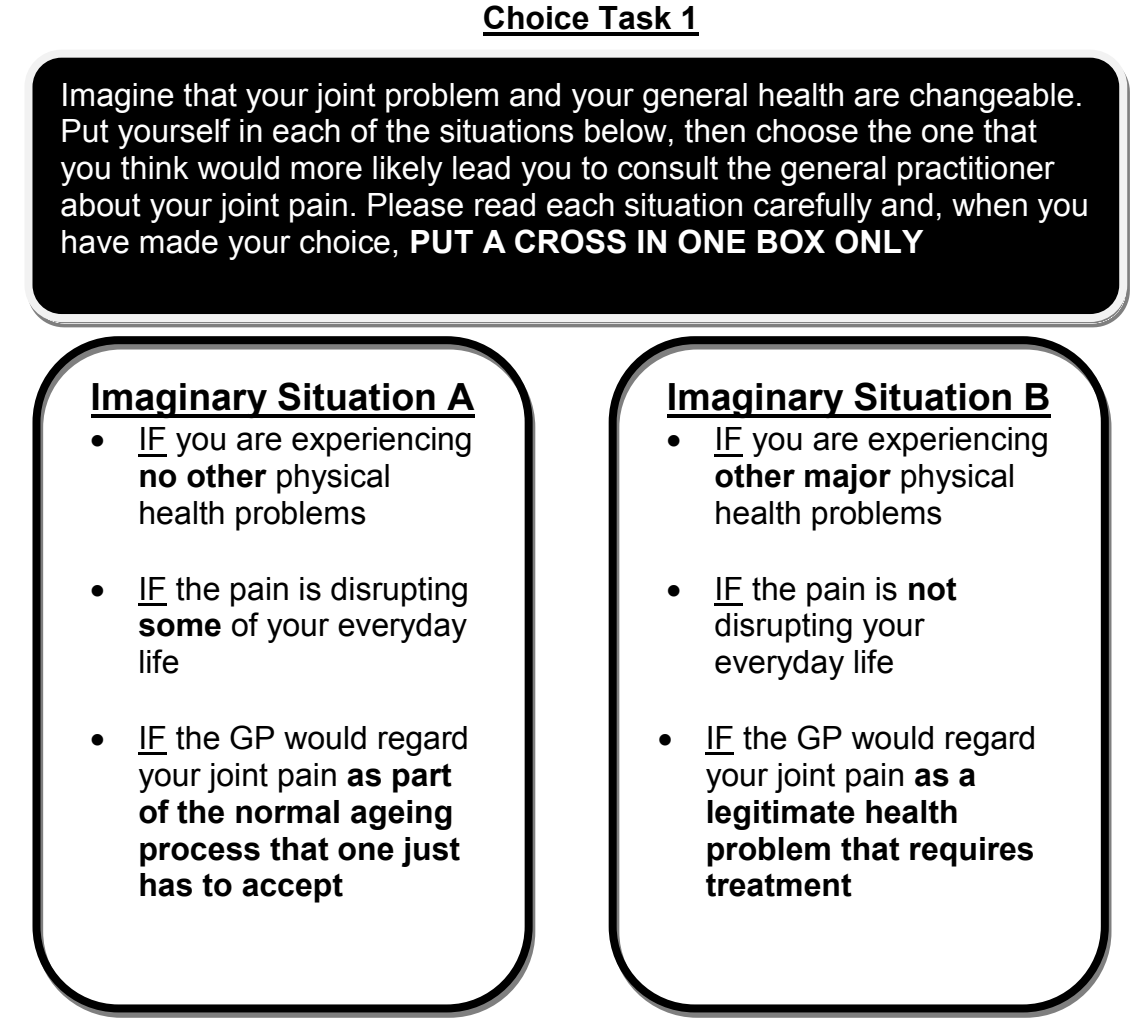
Respondents were asked to choose between 2 alternative scenarios (a pairwise choice set) and indicate under which scenario they would be more likely to go to the GP. Each scenario included 3 attribute levels and there were 10 different pairs of scenarios. These scenarios are made up of attributes presented at different levels. Feedback from the RUG identified that scenarios with more than 3-4 attribute levels to consider were likely to be cognitively demanding and burdensome (see chapter 4). A partial-profile design was therefore preferred to full-profile design. In partial-profile designs scenarios only include a subset of the total number of attributes that are being used. The main design included 10 random choice tasks (excluding holdouts) and 10 questionnaire versions (see section 5.6.4).

Table 6.1: Content of the survey instrument for main study

Section		Pages	General content	Format
Instructions		2	-	-
A	Factors that may influence your decision to consult	2	Direct rating of attribute importance	11-point NRS for all 6 attributes
B	Choice tasks	11	Choice tasks	Partial profile choice-based conjoint comprising 10 pairwise choice tasks with no holdouts
			Feedback on choice tasks	Single item on difficulty of completion Two open-ended questions with free-text responses
C	Joint pain	4	Joint pain: history, healthcare use, illness perceptions	†Presence of hip, knee, hand pain in past 12 months
				‡Episode duration - single item
				Time since onset - single item
				¶Consultations with healthcare professionals in past 12 months - GP, hospital specialist, physiotherapist, OT, nurse, chiropractor, osteopath, acupuncturist, homeopath, aromatherapist, other
				Investigations for joint pain in past 12 months - XRay, MRI, other
				Time since first visit to GP - single item
				§Illness perceptions - 9-item BIPQ
D	About you	1	Basic sociodemographic data	Date of Birth, Gender, Current employment status
E	Continuing to help with this study	2	Consent to further contact	Consent form, contact details of respondent and PI

GP General practitioner, OT Occupational therapist, MRI Magnetic resonance imaging, BIPQ Brief Illness Perceptions Questionnaire
† Adapted from Jinks et al. (2004)
‡ Adapted from de Vet et al. (2002)
¶ Adapted from Jinks et al. (2004)
§ (Broadbent et al. 2006)
The full questionnaire is reproduced in Appendix 8a (page 399)

Figure 6.1: Structure of the PPCBC questionnaire



I would be more likely to consult the doctor about the joint problem if I was in…

Situation A	Situation B
☐	☐

6.3.6 Databases and data handling

Mailing databases

Mailing databases were designed by specialist staff at the Arthritis Research UK Primary Care Centre at Keele University, using the same format as successfully used for NorStOP.

Data entry

The study team determined coding of questionnaires prior to data entry. Each patient had a unique anonymised code.

Data were manually entered into a database specifically designed for this study. Prior to data entry, this database was tested using a set of dummy data. Data were entered by experienced members of the administration team. Data entry databases are separate from mailing databases. All databases are secure and password-protected and conform to current data protection laws.

Data cleaning

Once all the data was entered the data gathered was cleaned under the supervision of the study statistician.

Data storage

Completed questionnaires were stored at the Arthritis Research UK Primary Care Centre without identifiable names and addresses and will be kept for twenty years in accordance with the Medical Research Council guidelines. Completed consent forms were securely stored at ARUKPCC separately from the questionnaires.

6.3.7 Data protection and confidentiality

Confidentiality issues

In line with the ARUKPCC Research and Governance Framework all personal identifiable information and details about patients and family members were kept separately from the research database during the mailing period. Participants were identified by the study codes. Only the research team has access to the research data and this information will also be kept on the ARUKPCC's central network drive. No information on patient's details or research data were stored on personal computer hard drives, laptops, disks or other means where data could be transferred. Future linkage of the study codes to personal identifiable data could only be achieved through re-contacting the patient's GP after the end of the study mailing period. Participants contact details were erased 3 months after final mailing.

There were secure physical storage arrangements for the hard copy data at the ARUKPCC within lockable filing cabinets. In addition, any hard copy research data that had been printed for checking were destroyed by shredding. The ARUKPCC also operates a key code entry system to ensure only appropriate persons have access to the building.

Breaching confidentiality

It was not anticipated that a situation would arise from this questionnaire study where the confidentiality of a participant would be breached. However, if there were such information within a questionnaire that suggested that the person were to harm him or herself or another person their GP would be contacted in the first instance.

6.4 Statistical analysis

6.4.1 Target sample size

In the systematic review response rates ranged from 18% to 94%, with a median of 60.5% and the majority ranging between 20 and 60%. Assuming a response rate of 40%, a sample size of 1000 is required to obtain data from 400 respondents. Conjoint analysis studies typically have between 150 and 1200 respondents (Orme 2010).

There is no accepted formula for determining minimum sample sizes for choice-based conjoint studies. However, one suggested rule of thumb is based on results from marketing studies (Johnson & Orme, 1996). Johnson and Orme (1996) suggest that the number of observations available to estimate each parameter should be ≥ 500 whereby: (nta/c ≥ = 500), where n=number of respondents, t=number of tasks, a=number of alternatives per task, and c=maximum number of levels per attribute. 500 is intended to be the minimum threshold and researchers can assume the stability of estimates when each main effect (attribute level) of interest is represented across the design around 500 times (Johnson & Orme, 1996).

Thus, assuming a minimum threshhold of 500 representations per main-effects level, in a full CBC profile the minimum number of respondents would be: (X respondents) = 500* (3 Levels) / [(10 tasks)*(2 alternatives)] = 75 respondents (Johnson & Orme 1996). However, in a partial-profile design the probability of each attribute being shown in any task is 50% (3 attributes/6 attributes in total). Consequently, the formula becomes (X respondents) =500* (3 levels)/[(10 tasks) (0.5 the probability that attribute is shown)*(2 alternatives)] = 150 respondents.

Thus a partial-profile design would yield a minimum sample size of 150 (Johnson & Orme 1996).

In a previous conjoint analysis study in osteoarthritis there was a sample size of 100, although this used an adaptive computerised form of conjoint analysis (Fraenkel et al. 2004a). However, sample sizes of 300 are generally recommended for estimating main-effects in full-profile choice-based conjoint (CBC) designs (Orme 2010) and partial-profile designs typically require slightly larger sample sizes still, hence the target of 400.

6.4.2 Planned data analysis

Throughout this thesis the significance level was set at 0.05.

The primary planned analyses are briefly summarised below.

Descriptive characteristics of respondents and non-respondents

Objective: To understand the descriptive characteristics of respondents and to identify possible selection bias due to non-response.

Data: Demographic, socioeconomic, and health characteristics of respondents obtained from NorStOP 6YFU and Section D of the survey questionnaire.

Analysis*:* Descriptive statistics (mean, SD; median, inter-quartile range; frequencies and percentages) and simple parametric and non-parametric methods for testing the statistical significance of differences between respondents and non-respondents (SPSS Inc, Chicago, IL).

Description of missing data and patterns of response to choice tasks

Objective: To indirectly gauge the comprehensibility and respondent burden of the choice tasks by examining the rates and patterns of missing data.

Data: Data from Section B of the survey questionnaire; all respondents

Analysis: Simple descriptive statistics (frequencies and percentages).

Counts Analysis

Objective: To describe the number of times each attribute level was chosen in the choice tasks.

Data: Responses from all respondents for a total of 6 attributes covering 16 levels.

Analysis: Sawtooth Software Market Research Tools (7.0) (SMRT) was used to conduct counts analysis - a rough method for summarising preferences for a sample. It calculates a proportion for each attribute level that is based on the number of times a level is chosen divided by the number of times it was available (presented to respondents). It shows how often a level was chosen when it was available. The proportion is a ratio when compared within the same attribute, though because preference for an attribute level depends upon desirability of the other levels within same attribute, it is not appropriate to directly compare a count proportion from one attribute level to a level from a different attribute (Orme 2010).

Standardised partworth utilities for each attribute level

Objective: To calculate the standardised (partworth) utility values for each attribute level using multinomial logit analysis (MNL).

Data: Responses from all respondents for a total of 6 attributes covering 16 levels.

Analysis: Sawtooth Software's CBC multinomial logit (MNL) programme was used to generate standardised partworth utility means and standard errors for each of the attribute levels.

Sawtooth Software's CBC package includes a multinomial logit (MNL) module. MNL is a multivariate statistical model which combines all respondents'

responses and estimates a single set of effects (partworth utilities) based on the total sample (the aggregated model).

These utilities can then be used when two or more alternatives/scenarios are available to estimate the probability of selecting each scenario based on that scenario's attribute levels. This is explained in greater detail in Appendix 9.

Relative importance of attributes

Objective: To calculate the relative importance of each attribute using multinomial logit standardised (partworth) utility values for each attribute level.

Data: Standardised partworth utilities for each attribute level.

Analysis: Relative attribute importance is calculated by finding the percentage of the range in utilities (maximum minus minimum utility) across attributes. This provides an intuitive measure of importance. However, without reference to the specific attribute levels importance this is meaningless because they are linked to the attribute level ranges that are used in the experiment (Orme 2006).

The relative utility of specific clinical scenarios

Objective: To construct hypothetical scenarios and use the overall utility of these scenarios to predict which of two scenarios is more likely to lead to GP consultation.

Data: Standardised partworth utilities for each attribute level.

Analysis: Referred to in the marketing literature as 'share of preference', this refers to the amount of respondent desirability captured by product alternatives in a market simulation, expressed as percentages summing to 100% across competing alternatives. Share of preference represents 't*he percentage of*

respondents projected to choose an alternative' or '*the average probability of respondents in the sample choosing an alternative'* (Orme 2006, p.151).

In the context of this PhD hypothetical scenarios were constructed by specifying levels on selected attributes and setting all other attribute levels as constant. The overall utility of these scenarios can be used in order to predict which of the two scenario is more likely to lead to GP consultation. The partworth utilities are then used to estimate strengths of preference for each scenario, and results are accumulated over respondents to provide shares of preference among scenarios (Orme 2010).

The relative propensity to consult can be discerned by adding the partworth utilities of the component attribute levels. For example, consider the two hypothetical scenarios in Table 6.2.

Table 6.2: Example of estimated strengths of preference for two hypothetical scenarios

	Scenario A		Partworth	Scenario B		Partworth
Attribute level	2.3	Pain is disrupting most everyday life	0.45	2.2	Pain is disrupting some everyday life	0.20
Attribute level	6.1	GP regards joint pain as part of the normal ageing process that one just has to accept	-0.43	6.2	GP regards joint pain as a legitimate health problem that requires treatment	0.43
Scenario total	0.02			0.63		
Exponentiated total	1.02 (/2.90)			1.88 (/2. 90)		
Share of preference (%)	35%			65%		

This shows that (when all other attribute levels are fixed as constant across scenarios) 65% of respondents would be more likely to consult the GP for joint

pain if they were in scenario B than if they were in scenario A. Conversely 35% of respondents would be more likely consult the GP for joint pain under scenario A than under scenario B.

The share of preference model uses the logit rule for estimating shares. The scenario utilities are exponentiated and shares are normalised to sum to 100% (Orme 2010). These are ratio scaled.

Comparing direct rating of attribute importance with findings obtained from conjoint analysis

Objective 1: To determine whether the direct ratings (of importance of each attribute) differed significantly from the highest ranked direct rating (attribute 6: the GP attitude) in order to get an indication of whether respondents are discriminating between the attributes.

Objective 2: To compare the rank order of the direct ratings of the importance of attributes and the attribute importance estimates from conjoint analysis.

Data: Responses from all respondents for rating of importance of each attribute (Section A) and relative importance for all 6 attributes.

Analysis: A one-way repeated measures ANOVA was used to compare the means of the 6 direct ratings (of importance of each attribute). An *a priori* simple contrast test within the repeated measures ANOVA was then used to determine whether the other direct ratings (of importance of each attribute) differed significantly from the highest ranked direct rating (attribute 6: the GP attitude). If other ratings are significantly different from the highest ranked attribute, this suggests the respondents are overall clearly defining the most important attribute. This will be informative because if researchers are able to gain an understanding of whether respondents are already discriminating, it poses the question of

whether conjoint analysis offers anything more than this, i.e. conjoint analysis claims to obtain the relative importance of attributes by forcing respondents to make choices. Descriptive statistics (mean, SD; median, inter-quartile range) were generated for the direct rating of importance of each attribute and these were ranked in order from highest to lowest. The relative importance of the attributes was also ranked in order from highest to lowest. These ranks of the direct rating estimates and attribute importance estimates were then presented in a table and used to look at how these estimates differed from each other in terms of the six attributes included. This is consistent with previous studies that have compared conjoint estimates with traditional rating scales (Phillips et al. 2002a).

Subgroup analysis

Objective: To identify distinct groups of respondents (subgroups) based on the choice data (preferences) and compare them with respect to descriptive characteristics.

Data: Data from Sections A, B, C and D of the questionnaire (see Table 6.1 on page 177); 250 (randomly selected) respondents. Distinct groups of respondents were determined based on their responses to Section B. These subgroups were compared with respect to descriptive statistics (mean, SD; median, inter-quartile range)

Analysis: Latent Class Analysis (Sawtooth Software, version 7) was used to identify distinct subgroups of respondents with similar preferences within the choice data (Orme 2006). In this way it goes beyond the traditional analysis of CBC data (i.e. estimating average partworths for the whole sample of respondents) because it identifies unobserved subgroups. Latent Class allows subgroups to be identified based on the actual choices of individuals.

The probability of membership of each subgroup for 250 (randomly selected) respondents is determined and each respondent is allocated to their most likely subgroup (i.e. the one for which they have the highest probability of belonging) (Dunn et al. 2006). Multinomial logit was used to fit a set of partworth utilities to each of the subgroups (Orme 2006). For example, one subgroup may be characterised by respondents prioritising the GP attitude, whereas another might be characterised by respondents who prioritise the Primary care management options available. Partworth utilities, a main dependent measure in this study, reflect the relative influence of each attribute level on participant choices with higher utility values indicating a stronger preference. To estimate the relative importance of each attribute, the range of each attribute's utility was converted to a percentage of the sum of the utility value ranges of all attributes (Orme 2006).

The number of subgroup solutions was facilitated by the researcher specifying a specific range to investigate (e.g. from 2 to 5 subgroups). The criteria for selecting a subgroup solution are as follows: Goodness of fit statistics for each solution, the distinctiveness and meaningfulness of the subgroups (interpretability) and, for each subgroup, the average probability of membership for respondents allocated to that subgroup. After selecting the subgroup solution, the subgroups were then described and interpreted according to the attribute utility ranges and importance scores in the choice tasks. Categorical and continuous demographic variables can then be described using simple frequencies and compared across subgroups using chi-square and Kruskall-Wallis tests respectively (See page 252 for more detail).

6.5 Ethical Approval

A favourable ethical opinion for this study was received from Staffordshire NHS Research Ethics Committee on (10/11/10) (Reference: 10/H1203/63).

7 Chapter Seven: Results I: survey response and descriptive characteristics of respondents to conjoint study

7.1 Introduction

In order to fulfil objectives set out in section 6.2, a population-based sample frame was selected whereby suitable data had been collected (6-year follow-up of participants in NorStOP 1 and 2) and from which further data could be collected into the determinants of GP consultation using a PPCBC study. The primary aim of this study was to quantify the relative importance of selected clinical need and general practice service factors ('attributes') in the decision to consult the GP in older adults with joint pain/OA.

This chapter will present the flow of response to the conjoint analysis study, compare the respondents and non-respondents systematically based on descriptive data from the NorStOP 6-year follow-up questionnaire, present the frequency of missing data within returned questionnaires, and present the descriptive characteristics of the respondents to the conjoint analysis study.

A high survey response rate is important because it increases the precision of the survey estimates by reducing non-response. Non-response is defined as the *'failure to obtain observations on some elements selected for the sample'* (Kish 1965, p. 532). Common sources of non-response include: refusals to take part (due to ill health or finding the questionnaire difficult to complete) and returned blank questionnaires with no comments. Non-response can be random or selective. Selective non-response can be considered a systematic error because it occurs when there are systematic differences between those selected and not selected for study, i.e. specific groups in the source population are under (or over) represented in survey respondents. If these groups behave differently in terms of

the survey variables, estimates based on data from respondents only are likely to be systematically biased.

In previous published conjoint studies in the primary care setting, response rates have ranged from 18 (Scott & Vick 1999) to 94% (Gerard et al. 2008) with evidence of non-response from people who are younger and male (Cheraghi-Sohi et al. 2008). The approach for addressing a selective non-response bias taken in this section is to compare the demographic, socioeconomic, and health characteristics of respondents and non-respondents.

In addition to non-response from individuals, missing data within returned questionnaires can similarly result in systematic bias and loss of precision. The amount and pattern of missing data can provide some insight into items or whole sections of survey instruments that may be unduly burdensome, incomprehensible, or judged by respondents to be inapplicable or inappropriate. A broad distribution of descriptive characteristics within respondents and a good response rate will help to reduce selection bias (Martin, 2005).

7.2 Aims and objectives

Using data gathered from the main conjoint analysis study in community-dwelling adults aged 50 years and over with joint pain, the aim was to understand the descriptive characteristics of respondents, to identify possible selection bias due to non-response, and to indirectly gauge the comprehensibility and respondent burden of the choice tasks by examining the rates and patterns of missing data. Thus the objectives were:

1. To compare the respondents and non-respondents based on descriptive data.
2. To present the frequency of missing data within returned questionnaires.

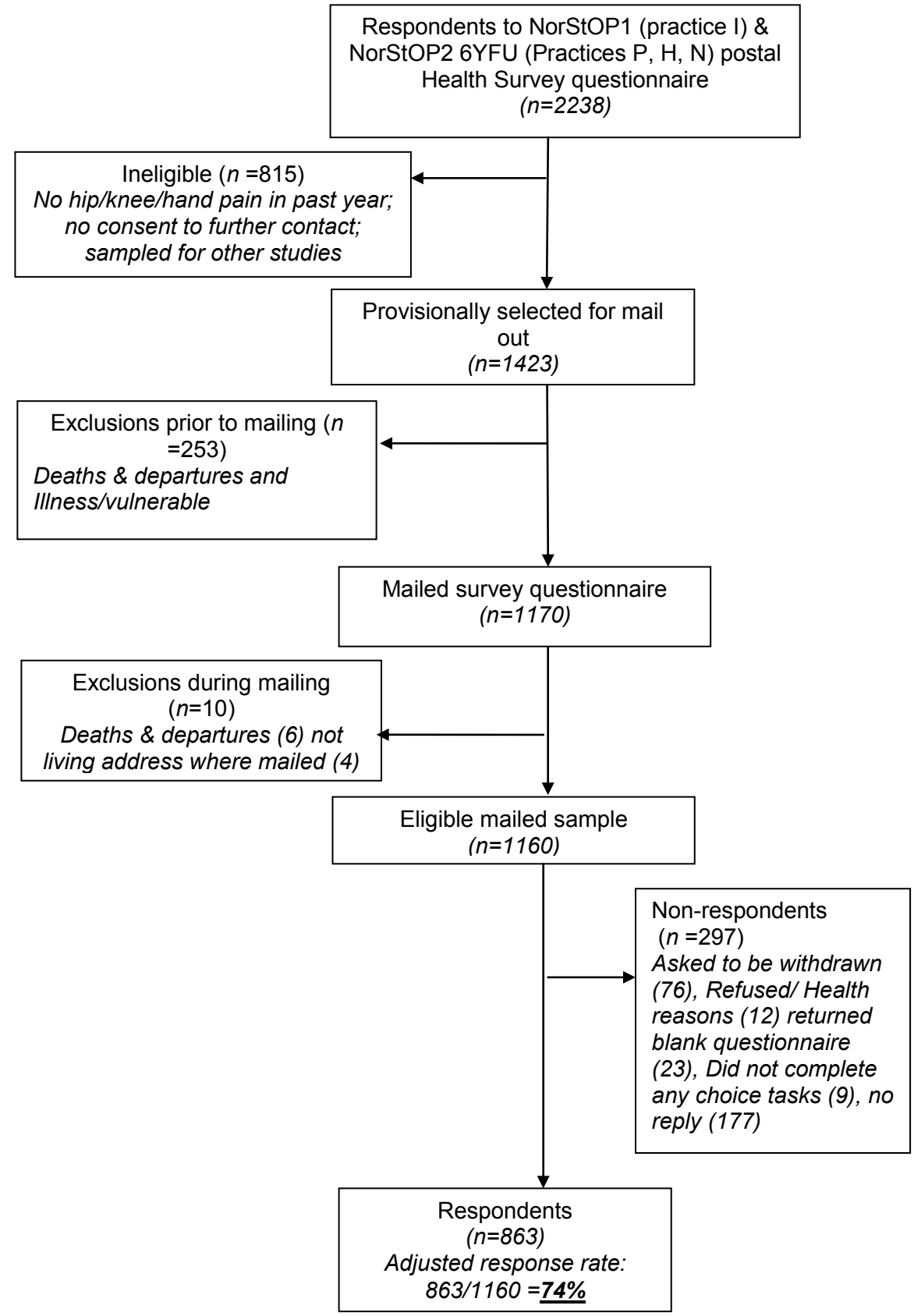
3. To present the descriptive characteristics of the respondents who comprise the participants in the conjoint analysis study.

The method of data collection and statistical analysis are presented in sections 6.3 and 6.4 respectively.

7.3 Response to the conjoint analysis study

Of the 2238 respondents to NorStOP1 and NorStOP2 6-year follow-up, 815 were ineligible due to not having hand, knee or hip pain in the past year, not consenting to further contact, or being sampled for other studies. A further 253 were removed due to deaths and departures, illness or vulnerability or being sampled for other studies. A further 10 were excluded during the mailing procedure when respondents had died or were not living at the address where mailed. Of the 1160 people selected to participate in the conjoint study, 863 (adjusted response rate: 74%) returned questionnaires (see Figure 7.1 overleaf)

Figure 7.1: Flowchart of response and participant flow



7.3.1 Respondents versus non-respondents

As might be expected, those who were excluded (recent deaths and those not living at the mailing address) were older.

In Table 7.1 (see page overleaf) χ^2 tests were used to compare respondents and non-respondents. An χ^2 test for trend was used for comparisons of age categories only. Table 7.1 shows that compared to non-respondents, respondents were slightly younger (χ^2 = 8.563, df=1, *p*=.003). There was no difference between non-respondents and respondents in terms of gender (χ^2 =.429, df=1, *p*=.512) or practice (χ^2 =1.952, df=3, *p*=.582), marital status (χ^2 =.242, df= 1, *p*=.623), employment status (χ^2 = 2.937, df=1, *p*=.087), living arrangement (χ^2 = .655, df=1, *p*=.418) or socioeconomic status (χ^2 = 1.933, df=2, *p*=.380).

Table 7.1: Demographic and socioeconomic characteristics of survey respondents, non-respondents, and exclusions

	Respondents	Non-respondents‡	Exclusions§
N	863	297	10
Age (years): mean (SD)	70 (7.5)	72 (8.3)	73.5 (8.5)
Age stratum (years):			
50-64	245 (28)	72 (24)	1 (10)
65-74	367 (43)	106 (36)	5 (50)
75+	251 (29)	119 (40)	4 (40)
Female	478 (55)	171 (58)	6 (60)
Practice:			
1	241 (28)	76 (26)	4 (40)
2	284 (33)	101 (34)	2 (20)
3	154 (18)	62 (21)	1 (10)
4	184 (21)	58 (20)	3 (30)
Married/cohabiting	627 (73)	214 (72)	6 (60)
Lives alone	185 (21)	71 (24)	5 (50)
Employment status:			
Full time paid	98 (11)	30 (10)	- (<1)
Part-time paid (not retired)	42 (5)	13 (4)	- (<1)
Part-time paid (part retired)	56 (6)	11 (4)	- (<1)
Sick <6months	2 (<1)	-	- (<1)
Sick >6months	12 (1)	6 (2)	1(10)
Fully retired, reached age	397 (46)	154 (52)	5(50)
Fully retired, early	96 (11)	34 (11)	- (<1)
Fully retired, ill health	75 (9)	15 (5)	1(10)
Carer	3 (<1)	4 (1)	- (<1)
Unemployed	5 (<1)	2 (1)	- (<1)
Look after home	21 (2)	8 (3)	- (<1)
Perceived financial strain†			
Strain	18 (2)	10 (3)	- (<1)
Be careful	305 (35)	110 (37)	2(20)
Little difficulty	359 (42)	133 (45)	3(30)
Comfortable	150 (17)	33 (11)	3(30)
Socioeconomic classification:			
Higher managerial	32 (4)	12 (4)	0(<1)
Higher professional	32 (4)	8 (3)	0(<1)
Lower managerial/professional	139 (16)	46 (15)	1(10)
Intermediate occupations	143 (17)	36 (12)	1(10)
Self-employed	47 (5)	18 (6)	1(10)
Lower supervisory/technical	67 (8)	32 (11)	1(10)
Semi-routine occupations	155 (18)	60 (20)	4(40)
Routine occupations	196 (23)	60 (20)	1(10)

Figures are numbers of participants (percentage) unless otherwise stated
Individual items may not add to totals due to missing data
‡ Includes those who did not return a questionnaire, returned a blank questionnaire and those who contacted the ARUKPCC to request they be removed from the study, but did not fulfil exclusion criteria
§ Includes those who died, were identified as being vulnerable or suffering from ill health or not living at the address where the questionnaire was mailed. These people were therefore unable to complete the questionnaire
† From (Thomas 1999)

Table 7.2 (see page overleaf) shows the distribution of general health characteristics of the sample broken down into survey respondents, non-respondents, and exclusions. Those who were excluded (included recent deaths and those who were no longer at the mailing address) had poorer self-rated health and were more likely to report all comorbidities included in the survey, with the exception of stroke and liver disease.

No difference in the frequency of number of self-reported comorbidities was found between non-respondents and respondents ($U(21) = 121409$, $Z = -1.362$, $p=.173$).

Table 7.2: General health characteristics of survey respondents, non-respondents, and exclusions

	Respondents	‡Non-respondents	§Exclusions
N	863	297	10
Self-rated health:			
Excellent	31 (4)	10 (3)	- (<1)
Very good	208 (24)	57 (19)	1 (10)
Good	395 (46)	130 (44)	6 (60)
Fair	181 (21)	83 (28)	2 (20)
Poor	43 (5)	15 (5)	1 (10)
Self-reported comorbidity:			
High blood pressure	401 (46)	138 (46)	6 (60)
Breathlessness	337 (39)	120 (40)	3 (30)
Memory	288 (33)	103 (35)	4 (40)
Previous fracture: other	273 (32)	98 (33)	3 (30)
Swelling ankles/feet	253 (29)	98 (33)	4 (40)
Leg circulation problems	251 (29)	88 (30)	5 (50)
Weakness in arm/leg	214 (25)	79 (27)	3 (30)
Eyesight problems	202 (23)	71 (24)	4 (40)
Dizziness	186 (22)	85 (29)	5 (50)
Chest problems	187 (22)	74 (25)	3 (30)
Heart problems	177 (21)	79 (27)	3 (30)
Cough	188 (22)	62 (21)	1 (10)
Deafness	177 (21)	56 (19)	3 (30)
Falls	129 (15)	44 (15)	1 (10)
Diabetes	128 (15)	42 (14)	1 (10)
Previous fracture: wrist	116 (13)	50 (17)	3 (30)
Shaking in hands	65 (8)	27 (9)	- (<1)
Previous stroke	42 (5)	16 (5)	- (<1)
Cancer	28 (3)	13 (4)	1 (10)
Previous fracture: hip	20 (2)	12 (4)	1 (10)
Kidney disease	18 (2)	7 (2)	3 (30)
Liver disease	11 (1)	2 (1)	- (<1)
Number of self-reported comorbidities (0-22) median (IQR)	4 (2,6)	4 (2,7)	6 (2,9)
HADS (0-21): median (IQR)			
Anxiety	5 (2.5,8)	6 (3,8)	6 (1,7)
Depression	3 (1,6)	4 (1,7)	6 (2,7)

Figures are numbers of participants (percentage) unless otherwise stated
HAD Hospital Anxiety & Depression scale (Zigmond & Snaith 1983); IQR Inter-quartile Range
Individual items may not add to totals due to missing data
‡ Includes those who did not return a questionnaire, returned a blank questionnaire and those who contacted the ARUKPCC to request they be removed from the study, but did not fulfil exclusion criteria
§ Includes those who died, were identified as being vulnerable or suffering from ill health or not living at the address where the questionnaire was mailed. These people were therefore unable to complete the questionnaire

7.4 Completeness of the data

Table 7.3 (see page overleaf) shows the levels of missing data for the direct rating of importance of each attribute and for each of the 10 choice tasks among conjoint analysis study respondents. In general these sections of the questionnaire were well-completed. The rate of missing data was 4-6% for direct rating and 2-4% for individual choice tasks. There was no evidence of a pattern of missingness for the choice tasks (e.g. more missing data for later choice tasks that might indicate progressive respondent burden). Of the 863 respondents, 792 (92%) had complete data on all choice tasks and 71 (8%) had partial data on choice tasks. Those who returned questionnaires but failed to complete any of the choice tasks were treated as non-respondents (n=9). This left questionnaires from 863 respondents (data for 8457 choice tasks), a response rate of 74% (863/1160) for the CBC tasks.

Rates of missing data were less than 10% for all other covariates included in the conjoint analysis study questionnaire (see Table 7.3 overleaf). The median score for the difficulty of imagining the hypothetical scenarios (based on an ordinal scale of 1= not at all hard to 5= extremely hard) for the conjoint sample was 2 (IQR=1, 3) (see Table 7.3 overleaf).

Table 7.3: Rates of missing data for direct rating/choice tasks and difficulty rating of choice tasks among survey respondents (n=863)

	N (%)
Direct rating of importance:	
Attribute 1: Pain characteristics	26 (3)
Attribute 2: Pain disruption	26 (3)
Attribute 3: Competing comorbidity	36 (4)
Attribute 4: GP assessment/investigations	39 (5)
Attribute 5: Primary care management/services	34 (4)
Attribute 6: GP attitude	31 (4)
Choice tasks:	
Choice task 1	14 (2)
Choice task 2	20 (2)
Choice task 3	19 (2)
Choice task 4	27 (3)
Choice task 5	26 (3)
Choice task 6	17 (2)
Choice task 7	17 (2)
Choice task 8	11 (1)
Choice task 9	12 (1)
Choice task 10	10 (1)
† Difficulty of hypothetical scenarios (1-5 ordinal scale): Median (IQR)	2 (1,3)
IQR Inter-quartile range; † Difficulty was measured on a 1-5 ordinal scale where 1 = not at all hard to imagine hypothetical scenarios and 5 = extremely hard to imagine hypothetical scenarios	

7.5 Descriptive characteristics of conjoint analysis study respondents

7.5.1 Demographic and socioeconomic characteristics of respondents

There were roughly equal numbers of respondents from each of the four participating practices (see Table 7.1 on page 195). The mean age of respondents was 70 years (SD 7.5) with the smallest age stratum being 50-64 years (n=245, 28%) (Table 7.1). 478 (55%) of respondents were female. The majority were married or cohabiting (n=627 (73%)) with 185 (21%) currently living alone. 609 (71%) of respondents were unemployed, with just 198 (23%) currently in employment. In terms of financial strain, 509 (59%) reported having 'little difficulty' or being 'comfortably off', while 305 (35%) reported financial strain. In terms of socioeconomic classification, 196 (23%) spent the majority of their life in routine occupations, 155 (18%) in semi-routine occupations, 143 (17%) in intermediate occupations and 139 (16%) in lower managerial/professional. Only 32 (4%) had worked in higher managerial/professional positions.

7.5.2 General health characteristics of respondents

Table 7.2 (see page 197) presents the general health characteristics of respondents. 224 (26%) reported 'fair' or 'poor' self-rated health. Self-reported comorbidity was common with high blood pressure (46%), breathlessness (39%), memory problems (33%), previous 'other' fractures (32%), and problems with leg circulation and swelling ankles/feet (29%) being among the most common of those included in the questionnaire. A small percentage of respondents had major illnesses such as cancer (3%), liver disease (1%) or kidney disease (2%). 128 respondents (15%) had diabetes and 5% had suffered a stroke.

The median number of self-reported comorbidities score (0-22) of respondents was 4 (IQR: 2,6) – which was similar to the median number of self-reported comorbidities score (0-22) of non-respondents, 4 (IQR: 2,7). The median number of self-reported comorbidities score (0-22) of exclusions was higher, 6 (IQR: 2,9).

The median HADs anxiety score (0-21) of respondents was 5 (IQR: 2.5, 8) and the median HADs depression score (0-21) was 3 (IQR: 1, 6) suggesting low levels of anxiety and depressive symptoms in this population-based sample.

Table 7.4: Rates of missing data for covariates in the conjoint survey questionnaire among survey respondents (n=863)

	N (%)
Employment status	56 (6)
Marital status	13 (2)
Live alone	11 (1)
Hip pain in past 12 months	10 (1)
Knee pain in past 12 months	5 (1)
Hand pain in past 12 months	7 (1)
Duration of present episode	16 (2)
Time since problem onset	11 (1)
Time since first GP consultation for joint pain	18 (2)
Healthcare consultations for joint pain in past 12 months:	
GP	-
Hospital specialist	-
Physiotherapist	-
Occupational therapist	-
Nurse	-
Chiropractor	-
Osteopath	-
Acupuncturist	-
Homeopath	-
Aromatherapist	-
Other	-
Investigations for joint pain in past 12 months	
MRI scan	1 (<1)
XRay	-
Other	1 (<1)
Brief Illness Perception Questionnaire (BIPQ)†:	
BIPQ 1: Consequence	14 (2)
BIPQ 2: Timeline	35 (4)
BIPQ 3: Personal control	28 (3)
BIPQ 4: Treatment control	25 (3)
BIPQ 5: Identity	21 (2)
BIPQ 6: Emotional representation	17 (2)
BIPQ 7: Causal	72 (8)
BIPQ 8: Illness concern	19 (2)
BIPQ 9: Coherence	19 (2)

† Broadbent et al. 2006

7.5.3 Joint pain and healthcare utilisation characteristics of respondents

Of the three peripheral joints included in the eligibility criteria for this study, knee pain was the most common (73%) followed by hand pain (68%) and hip pain (56%) (see Table 7.5 on page 204). The high frequency of each confirms that many had multiple-site joint pain. There were 71 (8%) with hip pain only, 145 (17%) with knee pain only, 116 (13%) with hand pain only. There were 61 (7%) respondents with hand + hip, 145 (17%) with hand + knee, 99 (11%) with hip + knee and 219 (25%) with pain (hand + hip + knee) in all three sites in the past 12 months.

Most respondents had longstanding joint problems with only 32 (4%) reporting the onset of their problem in the past year. 336 (39%) reported their joint problem first beginning 10 or more years ago and 275 (32%) reported first consulting their GP for this over 10 years ago. Despite the long history of joint problems, many (38%) reported experiencing a month free of joint pain within the previous 3 months suggesting that most respondents had a long history of intermittent, episodic joint pain. An estimated 50% of respondents had consulted their GP within the past 12 months about their joint pain. 141 (16%) reported never having consulted the GP about their joint problem. Whilst the GP was the health professional most commonly consulted within the past 12 months about joint pain, 163 (19%) of respondents had consulted a hospital specialist, 148 (17%) had consulted a physiotherapist, and 67 (8%) had consulted a nurse in the same period. Consultation with other healthcare professionals and complementary/alternative therapists was less common (1-4% each). 204 (24%) of respondents had received an XRay in the past 12 months.

Table 7.5: Joint pain and healthcare utilisation characteristics among conjoint analysis study respondents (n=863)

	N (%)
Hip pain in past 12 months	483 (56)
Knee pain in past 12 months	633 (73)
Hand pain in past 12 months	589 (68)
Pattern of joint pain involvement:	
Hip only	71 (8)
Knee only	145 (17)
Hand only	116 (13)
Hand + Hip	61 (7)
Hand + Knee	145 (17)
Hip + Knee	99 (11)
Hand + Hip + Knee	219 (25)
Average pain intensity (0-10) Median (IQR)	
Hip pain intensity	1 (0, 5)
Knee pain intensity	3 (0, 7)
Hand pain intensity	2 (0, 5)
How long since pain first started:	
Less than 1 year ago	32 (4)
1-5 years ago	243 (28)
6-10 years ago	241 (28)
More than 10 years ago	336 (39)
†How long since month without joint pain:	
Less than 3 months	325 (38)
3-6 months	85 (10)
7-12 months	53 (6)
1-2 years	72 (8)
3-5 years	97 (11)
6-10 years	97 (11)
More than 10 years	118 (14)
When first consulted GP for joint pain	
Never been to GP	141 (16)
Less than 1 year ago	60 (7)
1-5 years ago	212 (25)
6-10 years ago	157 (18)
More than 10 years ago	275 (32)
Healthcare consultations for joint pain in past 12 months:	
GP	434 (50)
Hospital specialist	163 (19)
Physiotherapist	148 (17)
Nurse	67 (8)
Osteopath	36 (4)
Acupuncturist	33 (4)
Occupational therapist	23 (3)
Chiropractor	21 (2)
Aromatherapist	17 (2)
Homeopath	9 (1)
Investigations for joint pain in past 12 months:	
XRay	204 (24)
MRI scan	53 (6)
Other	27 (3)
MRI Magnetic resonance imaging † Adapted from (de Vet et al. 2002)	

7.5.4 Illness perceptions of respondents

Respondents scored highly (median 10; IQR 8, 10) on the Brief Illness Perceptions Questionnaire (BIPQ) item related to perceived timeline, indicating that most expected their joint pain to continue for a long time into the future (Table 7.6 on page overleaf). In contrast, the median score for BIPQ item 6 (emotional representation) was low (median 2; IQR 0, 6) suggesting that most respondents did not feel that their joint pain had a significant emotional impact on them. The scores for the remaining BIPQ items were generally distributed around a median of 5.

Responses to the causal attribution item 7 can be grouped into 11 categories in order for categorical analysis to be performed. The approach taken was to use the classification adopted in an earlier ARUKPCC study of knee osteoarthritis and hand osteoarthritis. These include: predispositions/heredity/mal-alignment/gender, natural degeneration/natural lifelong progression, ageing, environmental conditions, trauma/injury, specific pathology i.e. physiological disease/knee disorders/referred conditions/medical interventions, lifestyle (family/leisure i.e. overuse/misuse), lifestyle (occupation i.e. overuse/misuse), don't know and unclassifiable (Peat & Thomas 2009). Respondents attributed causality of their joint pain to lifestyle (occupation i.e. overuse/misuse) factors most frequently (21%), followed by ageing (17%) and specific pathology (14%).

Table 7.6: Brief Illness Perceptions Questionnaire scores and Coping Strategies Questionnaire among conjoint analysis study respondents (n=863)

	N (%)
† Brief Illness Perceptions Questionnaire (BIPQ) individual items (0-10): Median (IQR)	
BIPQ 1: Consequence	5 (2, 7)
BIPQ 2: Timeline	10 (8,10)
BIPQ 3: Personal control	5 (3, 8)
BIPQ 4: Treatment control	5 (3, 8)
BIPQ 5: Identity	6 (3, 8)
BIPQ 6: Emotional representation	2 (0, 6)
‡BIPQ 7: §Causal attributions:	
Lifestyle (occupation i.e. overuse/misuse)	178(21)
Ageing	144(17)
Specific pathology: physiological disease/knee disorder/referred condition/medical intervention	120(14)
Lifestyle (family/leisure i.e. overuse/misuse)	115(13)
Trauma/injury	81 (9)
Natural degeneration/natural lifelong progression	78 (9)
Predispositions/heredity/mal-alignment/gender	30 (3)
Environmental conditions	19 (2)
Don't know	14 (2)
Unclassifiable	12 (1)
BIPQ 8: Illness concern	5 (3, 8)
BIPQ 9: Coherence	3 (1, 5)
¶ Coping strategies questionnaire (CSQ) individual items 1-7 (number of participants, %, scoring >0 on each item):	
CSQ 1: Distraction	336 (39)
CSQ 2: Reinterpreting pain sensation	333 (39)
CSQ 3: Catastrophising	329 (38)
CSQ 4: Ignoring pain	591 (68)
CSQ 5: Praying and hoping	301 (35)
CSQ 6: Coping self-statements	704 (82)
CSQ 7: Increased behavioural activities	583 (68)

‡Figures are numbers of participants (percentage) unless otherwise stated
Individual items may not add to totals due to missing data
† (Broadbent et al. 2006)
§ (Peat & Thomas 2009)
¶ (Jensen et al. 2003)

7.6 Discussion

7.6.1 Principal findings

The adjusted response rate to this conjoint analysis study was 74%. This was based within a sample frame of participants in an ongoing cohort study. The demographic, socioeconomic and health characteristics are heterogeneous with over 200 participants within certain strata (e.g. age, gender, previous GP consultation) to support future stratified analysis. There were no large amounts of missing data and comparisons of respondents and non-respondents suggest that non-response increases as age increases.

7.6.2 Interpreting the principal findings in the context of previous studies

The generalisability of research findings depends not only on the sample frame selected but also on the sample recruited (Martin, 2005). In this study generalisability of findings to the target population (i.e. 'older adults with joint pain') was limited due to a) the choice of sample frame and b) the evidence that there was possible selective response.

The adjusted response rate to this conjoint analysis study was 74%. It has been suggested that response rates of 75% and above can be considered good (Bowling, 2005). This response rate is most likely a result of recruiting from a selected sample (i.e. the sample was taken from well-disposed respondents to an ongoing cohort study). In comparison to the 15 previous conjoint studies reviewed (see chapter 3) this study achieved a good response rate, within a sample selected for convenience and to maximise response. In the systematic review in chapter 3 the median response rate was 60.5%. Ten studies used self-complete postal questionnaires (response rates ranging from 18 to 78% with a median of

53%). Although compared to face-to-face recruitment (5 studies with a median response rate of 76%), postal administration was lower, the range was still quite acceptable (18 to 78%). Although only two of the reviewed studies used highly selected samples, others studies that did not did recognise the benefits of using a selected sample. For example, Cheraghi-Sohi et al. (2008) achieved a 53% response rate, something which the authors concluded was reasonable given a complex questionnaire and an unselected group of patients (Cheraghi-Sohi et al. 2008). Using only 10 choice sets and ensuring respondent efficiency seemed to ensure a good response rate in this study. However, in one reviewed study, despite including many scenarios (24 choice sets) there was a good response rate of 88% (Haas 2005). This might be due to the fact that recruitment was carried out by a market research company. Similarly, a randomised trial by Coast et al. (2007) in a healthcare setting of a long and short DCE questionnaire (both including four attributes), indicated that the use of 16 scenarios obtained a very similar response rate (83%) to those obtained using only 8 scenarios (85%) (Coast & Horrocks 2007).

Similarly to the present study, two of the reviewed studies used a highly selected sample and had limited generalisability (Morgan et al. 2000, Longo et al. 2006). Morgan et al. (2000) selected a random sample of respondents to a previous postal survey of 25,090 randomly selected Sheffield residents who gave permission to further contact. They attained a comparable response rate to the present study (65%) (Morgan et al. 2000). Morgan et al. (2000) concluded that despite the sample being randomly selected the population from which the sample was chosen may not be representative of the population of Sheffield (the target population).

Longo et al. (2006) mailed a DCE to participants recruited to a previous randomised trial and attained a comparable response rate to this study (78%). Non-respondents and respondents were compared in this study, although patients were recruited with certain medical conditions in order to assess patients' utilities in parallel with a randomised trial in training GPs in shared decision-making and risk communication skills. It would therefore not be appropriate to generalise these findings beyond that studies target population (Longo et al. 2006).

Moreover, as it appears that older adults may be under represented in the present study sample it is even less likely to be generalisable to the target population. Of the 15 studies reviewed in Chapter 3 (see Table 3.3 on page 69) four reported non-response (Cheraghi-Sohi et al. 2008, Longo et al. 2006, Gerard et al. 2008, Ryan et al. 1998), with one presenting a table of non-respondents characteristics (Longo et al. 2006) and one referring to results in another paper (Ryan et al. 1998). For Cheraghi-sohi et al. (2008) the only available data on non-respondents were their age and sex, and these data showed some bias in that older patients and women were more likely to respond. Conversely, in the present study older people were less likely to respond. There were no explanations given for similarities and differences between respondents and non-respondents, although the authors acknowledged that because results were based on responses from primary care patients in the UK, they may not be generalisable to patients in different settings. In the Gerard et al. (2008) study respondents were more likely to be women, to not care for someone at home, to use a car to get to the surgery, to be unemployed and to be attending the surgery for an ongoing medical problem. In the present study unemployment was not significantly associated with non-response. No explanations were given for similarities and

differences between respondents and non-respondents in this study. In the Longo et al. (2006) study respondents and non-respondents were described as comparable, since there were no significant differences between them. There were no further explanations given for similarities and differences between them. Overall, in the studies that did report non-response, little attention was given into the aspects of non-response bias or selective non-response.

Of the 15 studies reviewed in Chapter 3 (see Table 3.3 on page 69) only 3 of them reported missing data (Cheraghi-Sohi et al. 2008, Scott et al. 2003, Turner et al. 2007). Cheraghi-sohi et al. (2008) reported a small amount of missing data for characteristics (e.g. age, income, and long-term illness status). Scott et al. (2003) reported that data on the dependent and independent variables were missing (8%) (Scott et al. 2003). These rates of missing data are comparable to the present study, as overall missing data rates (that were reported) were not substantial. It appears as though the present study was consistent with missing data rates in previous studies using conjoint analysis.

7.6.3 Strengths and weaknesses of the study

The main limitation was that generalisability to all older adults with joint pain was not justified, despite the sample being comprised of a diverse range of characteristics. This was because the sample frame was comprised of participants in a recent local population cohort with joint pain, was selected for convenience and to maximise response. However, being cautious and recruiting from a sample 'well-disposed to research' - one used to returning questionnaires to ARUKPCC – ensured a good response rate.

Across the choice sets 1 to 10, there is no evidence of missing data being progressive (see Table 7.3 on page 199). The median score for the difficulty of

imagining the hypothetical scenarios (based on an ordinal scale of 1= not at all hard to 5= extremely hard) was low (2: 1,3). This may suggest that there was a fair level of acceptability of the PPCBC task to respondents because respondent burden did not increase dramatically as the questionnaire progressed. This may be a result of the number of criteria for optimal design of conjoint analysis studies outlined in the ISPOR checklist (Bridges et al. 2011) that were met in the design of this questionnaire (see chapters 4 and 5). These included the use of partial (based on 3 attributes) rather than full-profiles, the presentation of only two profiles per choice task, and an upper limit of 10 choice tasks feasible for completion. Ensuring these were optimised in the conjoint design may account for the high response rate (i.e. increased respondent efficiency) and the acceptability of the PPCBC to respondents. The main disadvantage of partial-profiling is that interactions cannot be tested for. However, this was decided to be an acceptable price to pay for the benefit of partial-profile design in reducing respondent burden.

A number of strategies were used that have been shown to increase response to postal surveys (Edwards et al. 2002). These include ensuring that the questionnaires were user-friendly to increase respondent efficiency and reminder postcards and repeat questionnaires were mailed to non-respondents. No incentives were offered, and all mailed respondents were offered the opportunity to opt-out of the study.

7.6.4 Conclusion

The combination of a 74% response rate and little missing data implies that this research design may be well-completed and acceptable to a majority of the target population of community-dwelling older adults with joint pain. Generally, response rates of 75% and above can be considered good (Bowling, 2005).

Similarly to previous conjoint studies (Longo et al. 2006) the partial-profile design used in this study achieved an acceptable response rate when administered within an existing cohort, one that has been primed and is ‘well-disposed to’ research.

8 Chapter Eight: Results II: direct rating of selected attributes, conjoint utilities and relative importance of attributes

8.1 Introduction

Chapter 7 presented the flow of response (response rate of 74%), compared respondents and non-respondents based on descriptive data and presented the descriptive characteristics of the respondents to the conjoint analysis study. In this chapter the focus is on presenting the relative importance of selected determinants (attributes) of the decision to consult the GP about joint pain (section 8.3), how these attributes may be traded-off in the context of different clinically relevant scenarios (section 8.4) and how conjoint estimates compare and contrast with the findings from the direct rating of attributes (section 8.5).

8.2 Aims

Using data gathered from the main PPCBC in community-dwelling adults aged 50 years and over with joint pain, the aims of the present analyses were:

1. To identify the relative importance of selected determinants (attributes) of the decision to consult the GP about joint pain.
2. To explore how attributes may be traded-off in the context of different clinically relevant scenarios.
3. To compare and contrast the findings obtained from PPCBC with those obtained from direct rating by respondents.

The method of data collection and statistical analysis are presented in section 6.4.2 (on page 182).

8.3 Relative importance of selected attributes

8.3.1 Conjoint analysis

Counts analysis

Objective: To describe the number of times each attribute level was chosen in the choice tasks.

Table 8.1 presents the simple frequency with which each attribute level was chosen in the choice tasks, expressed as a proportion of the number of times that it was presented. This provides a sense of whether respondents' perceived importance of levels within each attribute was in the direction expected as well as the basis for an initial crude estimate of the relative importance of each of the attribute levels from the choice tasks.

Assuming the frequency of choosing an attribute reflects its perceived importance, it can be seen that the pattern of responses was broadly in the direction expected, namely: more severe pain episodes, more disruption to everyday life, the availability of more thorough assessment from the GP, greater management options including the multidisciplinary team and novel new treatment, and a 'legitimising' GP attitude increased the probability of consulting the GP for joint pain. The exception was competing comorbidity. The presence of other major physical health problems increased the probability of consultation for joint pain; the opposite of what might be predicted if it was assumed that these other physical health problems would compete with joint pain to move it down the patient agenda.

A positive GP attitude (68%) and pain disrupting most of everyday life (65%) were the most frequently chosen attribute levels when presented in the choice tasks. However, these must clearly reflect in part the relative unimportance of the other levels within those attributes. Thus, a 'normal ageing/accept' GP

Table 8.1: Perceived importance of attributes from choice tasks: count analysis

Attributes and levels				Times presented	Times chosen	Selected (%)
1	Pain characteristic	1.1	Dull aching pain	4315	2040	47.3
		1.2	Severe unpredictable episodes	4315	2275	52.7
2	Level of disruption to everyday life	2.1	None	2840	835	29.4
		2.2	Some	2859	1590	55.6
		2.3	Most	2931	1892	64.5
3	Competing comorbidity	3.1	None	2944	1275	43.3
		3.2	Minor	2836	1390	49.0
		3.3	Major	2850	1651	57.9
4	GP assessment/ investigations	4.1	Subjective only	2853	1136	39.8
		4.2	Subjective-thorough physical exam	2842	1507	53.0
		4.3	Subjective-thorough physical exam-X rays/blood tests	2935	1678	57.1
5	Primary care management /treatment	5.1	Verbal advice-prescribed analgesia	2930	1178	40.2
		5.2	Written advice-prescribed analgesia-PN follow-up-physio referral	2858	1530	53.5
		5.3	Written advice-prescribed analgesia-PN follow-up-physio referral-promising new treatment	2842	1609	56.6
6	GP attitude	6.1	Normal ageing process-accept it	4315	1360	31.5
		6.2	Legitimate health problem-requires treatment	4315	2955	68.4

GP General practitioner, PN Practice Nurse
NB Total for all levels for each attribute equals 8630 = 863 respondents completing 10 choice tasks. Number of times each level was presented is effectively equal within each attribute confirming balanced design

attitude (32%) and absence of any disruption to everyday life (29%) can be interpreted as perceived disincentives to consulting the GP for joint pain.

For the three-level attributes (disruption with everyday life, competing comorbidity, GP assessment/investigations, and primary care management/treatment) the counts analysis also provides some insight into the magnitude of the difference between successive levels of each attribute. Thus, moving from 'none' to 'some' disruption is perceived as more important than the difference between 'some' and 'most'. Similarly, the difference between the most basic level of assessment/investigations and management/treatment and the next level (representing 'optimal' services within the scope of what is available and recommended) is much more important than the addition of extra investigations or promising new treatments.

The simple counts analysis provides a reasonable indication of the relative importance of the different attributes. The ratio of proportions in simple count analysis can approximate the relative importance of the attributes when using logits. However, they do not permit the modeling of trade-offs between attributes or the comparative utility of scenarios made up of combinations of attribute levels.

Multinomial logit: Standardised partworth utilities for each attribute level and attributes

Objective: To calculate the standardised (partworth) utility values for each attribute level and the relative importance of each attribute using multinomial logit.

Figure 8.1 illustrates the partworth utilities for all attribute levels for sample respondents as a whole. Note that the utilities are zero-centred within each attribute and thus the values are relative, not absolute (see Appendix 9 for more information). Thus, a negative value does not have an absolute meaning, for

example a negative value does not mean that people would actively stay away from the GP, but its meaning is relative to the other attribute levels (i.e. more/less likely to lead to consultation than the other attribute level(s).

As with counts analysis, the pattern of partworth utilities were broadly in the direction expected. Thus, the scenario with the highest utility (i.e. more likely to lead to GP consultation) will be one in which the person is experiencing joint pain that was disrupting most of their everyday life (0.45), if the GP would regard their joint pain as a legitimate health problem that requires treatment (0.43), if they had other major physical health problems (0.23), if the GP would investigate with appropriate x rays and blood tests as well as ask about their symptoms and their effect on their day-to-day life and conduct a thorough physical examination of the joints (0.21), if the GP would offer a promising new treatment as well as prescribe pain relief, give written advice about their condition and arrange follow-up with a practice nurse and physiotherapy referral (0.19) and if they were experiencing short episodes of more severe, often unpredictable pain (0.08).

A positive GP attitude (0.43) and pain disrupting most of everyday life (0.45) had the highest partworth. However, these clearly reflect partworth utilities of the other levels within those attributes. Thus, a ‘normal ageing/accept’ GP attitude (-0.43) and absence of any disruption to everyday life (-0.65) can be interpreted as perceived disincentives to consulting the GP for joint pain.

The partworth utilities reflect the findings from counts analysis of a large step in utility from ‘none’ to ‘some’ disruption and from ‘basic’ to ‘optimal’ assessment and management services.

In Table 8.2 the coefficients (partworth utilities) are presented again, together with an estimate of their precision. All standard errors were well below

0.05 (Orme 2010) implying acceptable precision. For more information on the statistical procedures used please refer to section 6.4.2 (on page 182).

Table 8.2 also presents the attribute utility ranges, which form the basis for quantifying the relative importance of the attributes. From these it is apparent that the decision to consult the GP for joint pain was most strongly influenced by the level of disruption to everyday life (relative importance of 31%). This was followed by GP attitude (24%). The importance of competing comorbidity (13%), primary care management/treatment (13%), and GP assessment/investigations (14%) were similar. Pain characteristics were least important (5%).

Figure 8.1: Zero-centred utilities for all 16 attribute levels

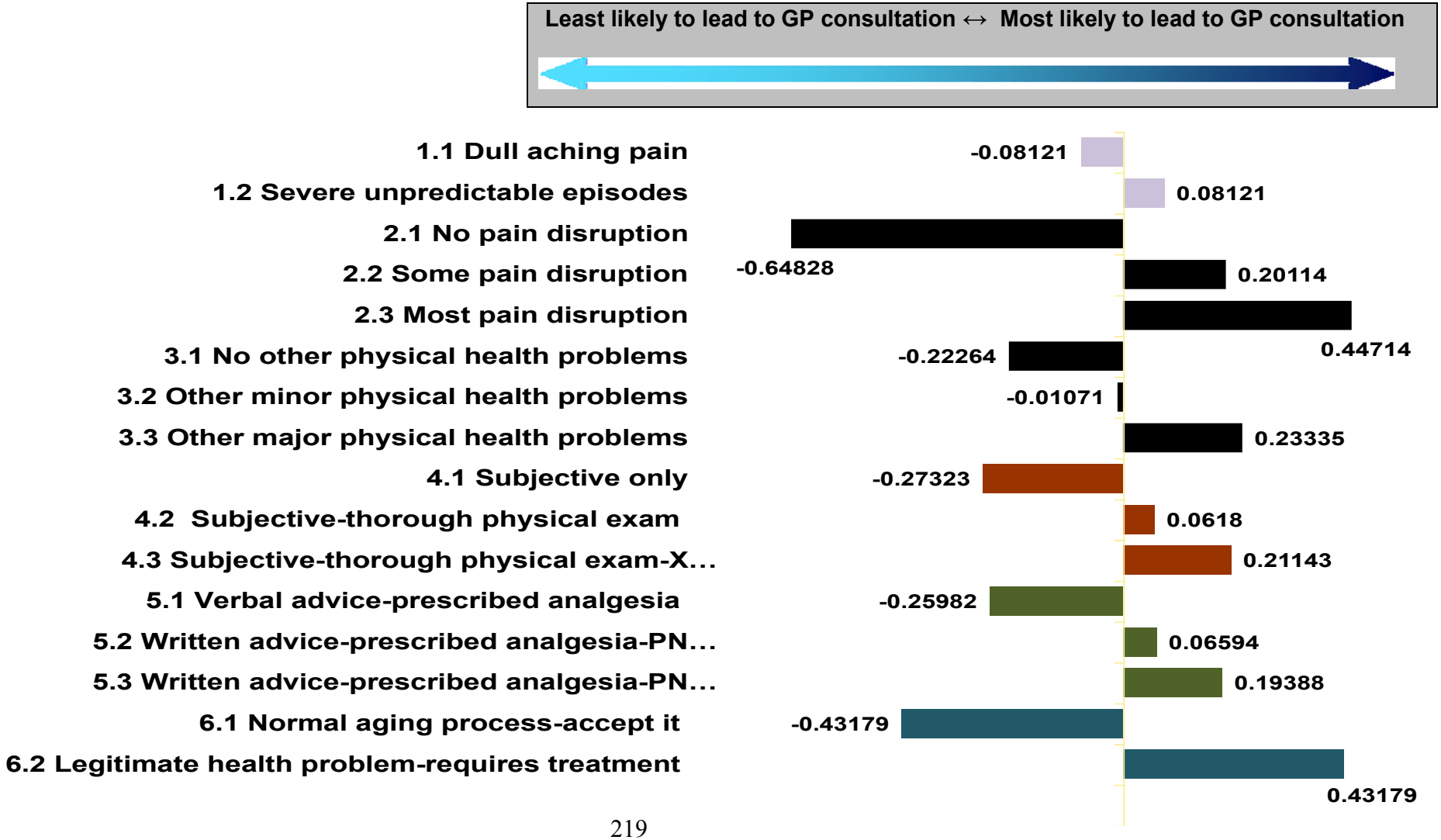


Table 8.2: Perceived importance of attributes and levels from choice tasks: multinomial logit

Attributes and levels				Coefficient (partworth utility)	SE	95%CI	Attribute utility range	†Attribute importance scores (%)	Rank
1	Pain characteristics	1.1	Dull aching pain	-0.08	0.017	(-0.114, -0.048)	0.16	5	6
		1.2	Severe unpredictable episodes	0.08	0.017	(0.048, 0.114)			
2	Level of disruption to everyday life	2.1	None	-0.65	0.030	(-0.706, -0.590)	1.10	31	1
		2.2	Some	0.20	0.028	(0.147, 0.255)			
		2.3	Most	0.45	0.028	(0.392, 0.502)			
3	Competing comorbidity	3.1	None	-0.22	0.027	(-0.276, -0.169)	0.46	13	4
		3.2	Minor	-0.01	0.028	(-0.065, 0.044)			
		3.3	Major	0.23	0.027	(0.179, 0.287)			
4	GP assessment/investigations	4.1	Subjective only	-0.27	0.028	(-0.328, -0.219)	0.48	14	3
		4.2	Subjective-thorough physical exam	0.06	0.028	(0.007, 0.117)			
		4.3	Subjective-thorough physical exam-X rays/blood tests	0.21	0.028	(0.157, 0.265)			
5	Primary care management/treatment	5.1	Verbal advice-prescribed analgesia	-0.26	0.027	(-0.312, -0.208)	0.45	13	5
		5.2	Written advice-prescribed analgesia-PN follow-up-physio referral	0.07	0.027	(0.012, 0.120)			
		5.3	Written advice-prescribed analgesia-PN follow-up-physio referral-promising new treatment	0.19	0.027	(0.141, 0.247)			
6	GP attitude	6.1	Normal ageing process-accept it	-0.43	0.018	(-0.466, -0.397)	0.86	24	2
		6.2	Legitimate health problem-requires treatment	0.43	0.018	(0.397, 0.466)			

CI confidence interval, GP General practitioner, PN Practice Nurse, SE Standard error, † = attribute utility range/sum total of attribute utility ranges

8.4 The relative utility of specific clinical scenarios

Objective: To construct hypothetical scenarios and use the overall utility of these scenarios to predict what scenario is more likely to lead to GP consultation.

All possible combinations of scenarios (2*3*3*3*3*2=324) were calculated and ranked in descending order of overall utility (see Appendix 10 for top 50 scenarios ranked in descending order). These can be used in order to predict what scenario is more likely to lead to GP consultation for joint pain. The higher the overall utility of the scenario is the greater the relative propensity to consult. The partworth utilities can be used to estimate strengths of preference for each scenario, and results are accumulated over respondents to provide shares of preference among scenarios (Orme 2010). The scenario utilities are exponentiated and shares are normalised to sum to 100% (Orme 2010). See section 6.4.2 (on page 182) for more information.

In this section, the focus is on pairwise analysis of selected clinical scenarios in response to the following questions:

1. To what extent might changing to a positive legitimising GP attitude precipitate the presentation of less disabling joint problems?
2. To what extent might changing to a positive legitimising GP attitude be more effective in encouraging consultations for joint problems than improving the rigour of assessment and the range of management options available?

8.4.1 To what extent might changing to a positive legitimising GP attitude precipitate the presentation of less disabling joint problems?

Table 8.3: Change in GP attitude vs change in pain disruption - some↔most

Attribute	Scenario A	Scenario B
2. Level of disruption to everyday life	Most (0.45)	Some (0.20)
6. GP attitude	Normal ageing, accept it (-0.43)	Legitimate health problem, requires treatment (0.43)
Total utility	0.02	0.63
†Preference	35%	65%
† Calculated as the exponentiated total utility/sum total of exponentiated utilities		

Table 8.3 shows that the difference between 'normal ageing-accept it' GP attitude and positive legitimising GP attitude is greater than the difference between some pain disruption and most disruption.

Table 8.3 shows that (when all other attribute levels are fixed as constant across scenarios) 65% of respondents would be more likely to consult the GP for joint pain if they were in scenario B than if they were in scenario A. Conversely 35% of respondents would be more likely consult the GP for joint pain under scenario A than under scenario B. This suggests that (all other things being equal) despite having 'most' pain disruption if respondents could expect a 'normal ageing-accept it' GP attitude then they would be less likely to consult than if they only had some pain disruption but they could expect the GP to have a legitimising attitude. Thus by changing the GP attitude to legitimising this would be estimated to have the effect of a net increase in consultations for people with 'some' disabling joint problems or at a time when disruption is less severe (i.e. earlier in the course of progression).

Table 8.4: Change in GP attitude vs change in pain disruption - none↔most

Attribute	Scenario A	Scenario B
2. Level of disruption to everyday life	Most (0.45)	None (-0.65)
6. GP attitude	Normal ageing, accept it (-0.43)	Legitimate health problem, requires treatment (0.43)
Total utility	0.02	-0.22
†Preference	56%	44%
† Calculated as the exponentiated total utility/sum total of exponentiated utilities		

Table 8.4 shows that the difference between 'normal ageing-accept it' GP attitude and positive legitimising GP attitude is not greater than the difference between no pain disruption and most disruption (with all else being equal). Table 8.4 shows that 56% of respondents would more likely consult the GP for joint pain if they were in scenario A than if they were in scenario B. Conversely 44% of respondents would be more likely consult the GP for joint pain under scenario B than under scenario A. This shows that the tendency to consult for non-disabling pain (given a positive legitimising GP attitude) would not reach the tendency to consult for severely disabling pain even when the patient expects a 'normal ageing-accept it' attitude from the GP.

8.4.2 To what extent might changing to a positive legitimising GP attitude be more effective in encouraging consultations for joint problems than improving the rigour of assessment and the range of management options available?

Table 8.5: Change in GP attitude vs improved management

Attribute	Scenario A	Scenario B
Primary care management/treatment	Written advice-prescribed analgesia-PN follow-up-physio referral-promising new treatment (0.19)	Verbal advice-prescribed analgesia (-0.26)
GP attitude	Normal ageing, accept it (-0.43)	Legitimate health problem, requires treatment (0.43)
Total utility	-0.24	0.17
†Preference	40%	60%
† Calculated as the exponentiated total utility/sum total of exponentiated utilities		

Table 8.5 shows to what extent a positive legitimising GP attitude has a greater influence upon the decision to consult than 'primary care management' (with all else being equal).

Table 8.5 shows that 60% of respondents would more likely consult the GP for joint pain if they were in scenario B than if they were in scenario A. Conversely, 40% of respondents would be more likely consult the GP for joint pain under scenario A than under scenario B. This suggests that respondents are willing to trade-off the offer of a promising new treatment, pain relief, written advice and follow-up with a practice nurse and physiotherapy referral in order to ensure the GP regards their joint pain as a legitimate problem.

Table 8.6: Change in GP attitude vs improved assessment and management

Attribute	Scenario A	Scenario B
Primary care management/treatment	Written advice-prescribed analgesia-PN follow-up-physio referral-promising new treatment (0.19)	Verbal advice-prescribed analgesia (-0.26)
GP assessment/investigations	Subjective-thorough physical exam-X rays/blood tests (0.21)	Subjective only (-0.27)
GP attitude	Normal ageing, accept it (-0.43)	Legitimate health problem, requires treatment (0.43)
Total utility	-0.03	-0.1
†Preference	52%	48%
† Calculated as the exponentiated total utility/sum total of exponentiated utilities		

Table 8.6 shows a tipping point (i.e. the point at which respondents are willing to trade-off one thing for another). In this comparison respondents are willing to trade-off a legitimating GP attitude (as in scenario B), in order to gain the offer of fully optimised primary care management/treatment and GP assessment/investigations packages (as in scenario A). 52% of respondents would more likely consult the GP for joint pain if they were in scenario A than scenario B. Conversely 48% of respondents would more likely consult the GP for joint pain under scenario B than under scenario A.

Table 8.7: Change in GP attitude vs improved assessment and management

Attribute	Scenario A	Scenario B
Primary care management/treatment	Verbal advice-prescribed analgesia (-0.26)	Verbal advice-prescribed analgesia (-0.26)
GP assessment/investigations	Subjective-thorough physical exam-X rays/blood tests (0.21)	Subjective only (-0.27)
GP attitude	Normal ageing, accept it (-0.43)	Legitimate health problem, requires treatment (0.43)
Total utility	-0.48	-0.1
†Preference	41%	59%
† Calculated as the exponentiated total utility/sum total of exponentiated utilities		

Table 8.7 looks at what aspects would realistically be within the GPs' power to change, i.e. highest level of GP assessment, lowest level of management.

Table 8.7 shows that 41% of respondents would more likely consult the GP for joint pain if they were in scenario A than if they were in scenario B. Conversely 59% of respondents would more likely consult the GP for joint pain under scenario B than under scenario A. Interestingly it appears that respondents are willing to trade the offer of fully optimised GP assessments/investigations to have their GP regard their joint pain as a legitimate problem. Table 8.8 shows to what extent a positive legitimising GP attitude and 'GP assessment' has a greater influence upon the decision to consult than 'Primary care management' (all else being equal).

Table 8.8: Change in GP attitude vs improved assessment and management

Attribute	Scenario A	Scenario B
Primary care management/treatment	Written advice-prescribed analgesia-PN follow-up-physio referral (0.07)	Verbal advice-prescribed analgesia (-0.26)
GP assessment/investigations	Subjective-thorough physical exam-X rays/blood tests (0.21)	Subjective only (-0.27)
GP attitude	Normal ageing, accept it (-0.43)	Legitimate health problem, requires treatment (0.43)
Total utility	-0.15	-0.1
†Preference	49%	51%
† Calculated as the exponentiated total utility/sum total of exponentiated utilities		

Table 8.8 looks at what aspects would be realistically modifiable within a wider multidisciplinary primary care team i.e. (GP, physiotherapist, nurse).

Table 8.8 shows that 49% of respondents would more likely consult the GP for joint pain if they were in scenario A than if they were in scenario B. Conversely 51% of respondents would more likely consult the GP for joint pain under scenario B than under scenario A. This shows that having their GP regard their joint pain as a legitimate problem is more important than receiving the offer of a fully optimised package of primary care management. Table 8.8 shows to what extent a positive legitimising GP attitude and 'GP assessment' has a greater influence upon the decision to consult than 'primary care management' (all else being equal).

8.5 Comparing direct rating of attribute importance with findings obtained from conjoint analysis

Objective 1: To determine whether the 6 direct ratings (of importance of each attribute) differed significantly from each other.

Objective 2: To compare the rank order of the direct ratings of the importance of attributes and the attribute importance estimates from conjoint analysis.

Respondents' direct ratings of the importance of each of the six selected attributes are shown in Table 8.9. Attributes related to service provision (GP attitude, primary care management options and assessment/investigation offered) had the highest mean importance rating among respondents. Competing comorbidity was rated the least important (but still with a mean rating of 4.66).

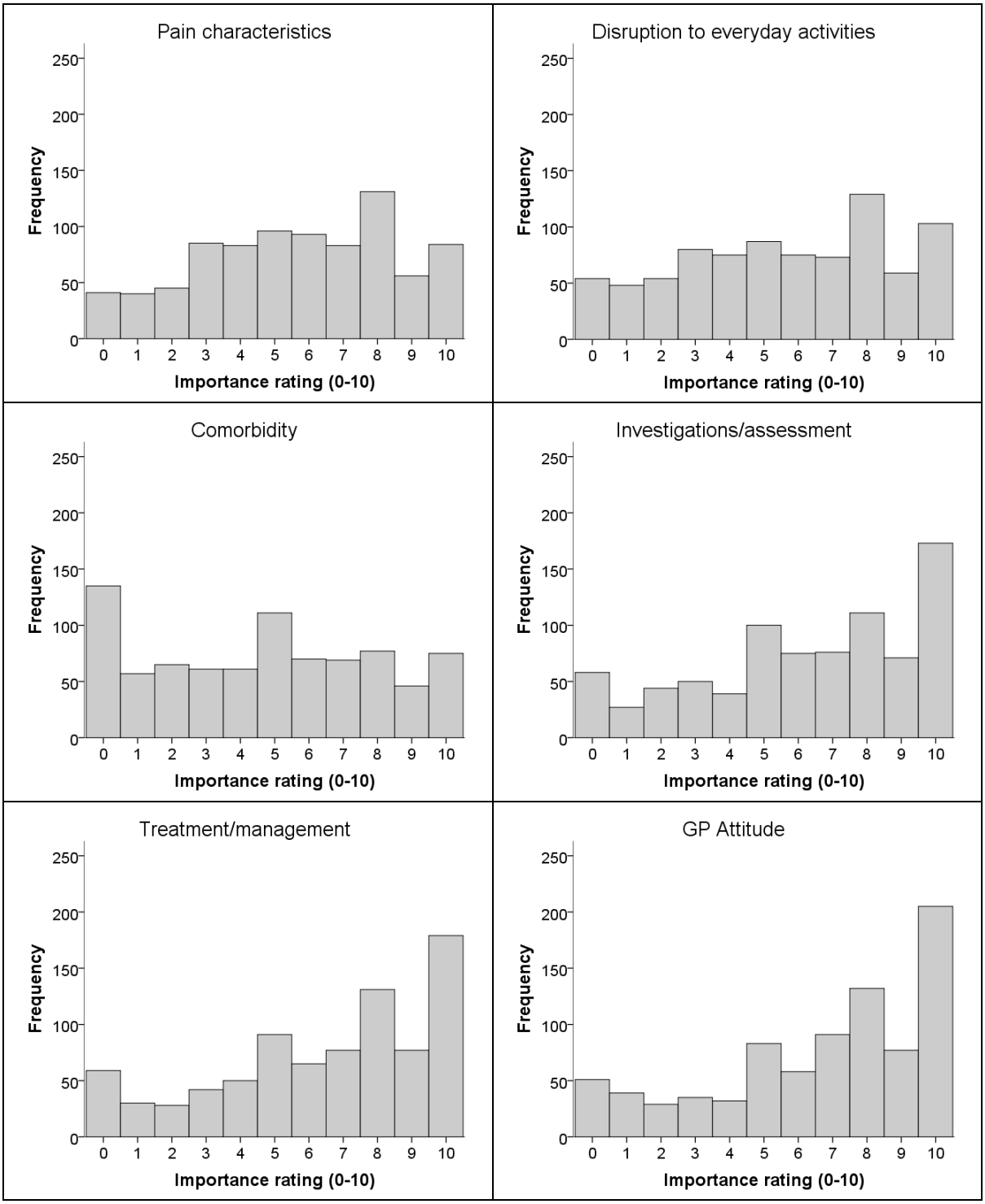
Table 8.9: Perceived importance of attributes from respondents' direct rating

Attribute	Median (IQR)	Mean (SD)	Rank
1. Pain characteristics	6 (3, 8)	5.65 (2.8)	4
2. Level of disruption to everyday life	6 (3, 8)	5.60 (3.0)	5
3. Competing comorbidity	5 (2, 7)	4.66 (3.2)	6
4. GP assessments/investigations	7 (4, 9)	6.26 (3.1)	3
5. Primary care management/treatment	7 (4, 9)	6.43 (3.1)	2
6. GP attitude	7 (5, 9)	6.65 (3.1)	1
IQR Inter-quartile range; SD Standard deviation Importance was rated on a 0-10 numerical rating scale where 0 = of no importance and 10 = extremely important			

A one-way repeated measures ANOVA with a Greenhouse-Geisser correction determined that the means of the 6 direct ratings (of importance of each attribute) differed significantly from each other (F (3.202, 2548.540) = 97.494, $p<.001$). There were significant differences between the mean scores of the five direct ratings when compared to the highest ranked mean, i.e. the GP attitude was

set as the reference category. The following were compared to GP attitude (mean=6.65, SD=3.1). i) pain characteristics (5.65 (2.8)) ; $F(1)$= 85.699, p<.001, ii) pain disruption (5.60 (3.0)); $F(1)$= 79.209, p<.001, iii) comorbidity (4.66 (3.2)); $F(1)$= 241.931, p<.001, iv) GP assessment/investigations (6.26 (3.1)); $F(1)$= 21.380, p<.001, v) primary care management (6.43 (3.1)); $F(1)$= 8.113, p=.005. Therefore, the other attributes appear to be significantly different from the highest ranked mean (GP attitude). This and the repeated measures ANOVA suggest that respondents are overall clearly defining the most important attribute.

Figure 8.2: Histograms of direct ratings of importance for the 6 attributes



From the histograms of the direct ratings of importance for the six attributes in Figure 8.2 (on page 230) it appears that there is evidence of ceiling effects (respondents rate attributes with the highest possible importance) and floor effects (respondents rate attributes with the lowest possible importance). In clinimetrics a rule of thumb often used is that there should not be more than 15% of the sample at floor or ceiling (Terwee et al. 2007). Thus compared with recommended guidelines there is evidence of a floor effect in the rating of attribute 3 (competing comorbidity = 16%) and a ceiling effect in attributes 4 (GP assessment= 20%), 5 (primary care management = 21%) and 6 (GP attitude= 24%).

Table 8.10: Compares rated importance of attributes and relative importance of attributes based on conjoint within respondents (n=863)

Attribute	Median (IQR)	Mean (SD)	Direct rating rank	Relative importance rank
1. Pain characteristics	6 (3, 8)	5.65 (2.8)	4	6 (5%)
2. Level of disruption to everyday life	6 (3, 8)	5.60 (3.0)	5	1 (31%)
3. Competing comorbidity	5 (2, 7)	4.66 (3.2)	6	4 (13%)
4. GP assessments/investigations	7 (4, 9)	6.26 (3.1)	3	3 (14%)
5. Primary care management/treatment	7 (4, 9)	6.43 (3.1)	2	5 (13%)
6. GP attitude	7 (5, 9)	6.65 (3.1)	1	2 (24%)
IQR Inter-quartile range; SD Standard deviation Importance was rated on a 0-10 numerical rating scale where 0 = of no importance and 10 = extremely important				

It was decided that the mean was more appropriate for ranking the direct rating of attributes and comparing subgroups, because the median inevitably leads to equal rankings. Table 8.10 shows that the direct rating of importance of

attributes mean scores were not consistent with the conjoint estimates. For example, the rating results (when placed in ranked order of mean value) indicate that the three service attributes (attributes 4, 5 and 6) are rated as more important than the three clinical need attributes (attributes 1,2 and 3). However, in the CBC model the partworth utilities of these attributes indicated that pain disruption (31%) was considerably more likely to lead to GP consultation for joint pain than GP assessments/investigations (14%) and primary care management/services (13%). The most striking difference between direct rating importance rankings and CBC rankings is for attribute 2, the level of disruption to everyday life. Attribute two became more important in the CBC exercise than it was in the direct rating.

8.6 Discussion

8.6.1 Principal findings

Although choice-based conjoint analysis is widely used by marketing researchers (Orme 2006), health economists (Ryan & Farrar 2000) and primary care researcher's (Cheraghi-Sohi et al. 2007) this is the only study to the researchers knowledge, to have applied these methods to the investigation of determinants of GP consultation for joint pain. Among the attributes studied, the level of disruption to everyday life caused by the joint pain emerged as the most important determinant of the decision to consult (31%), followed by the perceived attitude of the GP (24%), GP assessment (14%), primary care management (13%), competing comorbidities, (13%) and finally pain characteristics (5%).

In simulated clinical scenarios it was found that, in the most serious circumstances (most disruption to everyday life), patients will tend to consult the GP irrespective of their perceived attitude to joint problems. Changing the GP

attitude to positive and legitimising would not be expected to convert demand for consultation in non-disabling joint pain to the level seen in people with severely disabling pain who have low expectations/perceptions of their GPs attitude. However, compared to GPs with a known non-legitimising negative attitude to joint problems, GPs with a positive legitimising attitude will tend to see more people with joint problems due to the tendency of those with less severe disability to approach them. The only caveat to this is that if those GPs with a positive legitimising attitude offer only a very basic assessment and management service and the ones with a 'normal ageing-accept it' attitude offer a fully optimised service.

The GP attitude attribute (the attribute with the highest mean importance rating) appears to be significantly different from all other attributes. Therefore it might be interpreted that respondents are overall clearly defining the most important attribute. Evidence of floor and ceiling effects in some attributes reinforced the finding that respondents are already able to discriminate how important attributes are in the decision to consult.

The most prominent difference between the direct rating and CBC estimates (rankings) was for attribute 2, the level of disruption to everyday life. Thus, attribute two became more important in the CBC exercise than it was in the direct rating.

8.6.2 Interpreting the principal findings in the context of previous studies

The decision to consult the GP for joint pain was most strongly influenced by the disruption of pain to everyday life (relative importance of 31%). Its

importance was valued over six times more than pain characteristics (relative importance of 5%). This is consistent with the finding from a recent observational epidemiological study that disability is a greater determinant of help-seeking than pain severity for community dwelling adults with chronic pain in hip or knee (Thorstensson et al. 2009). Confirmation of this finding is useful (given some inconsistency in the observed data set out in chapter 1) but the current study goes further than this in quantifying the relative importance of 'disruption of pain to everyday life' compared to other attributes (such as attribute 1; pain characteristics). This study draws out that pain disruption is over six times more important than pain characteristics. This information is elicited specifically within the context of a dynamic internal process of respondents choosing scenarios, which would more likely lead them to consult the GP for joint pain.

Expecting to see a GP who regards their joint pain as a legitimate health problem was the second most important attribute (relative importance of 24%), valuing its importance almost twice as much as primary care management (relative importance of 13%) and competing comorbidities (relative importance of 13%). Results suggest that a change to positive legitimising GP attitude would be estimated to have the effect of a net increase in consultations for people with less disabling problems. However, it is important to note that the effect of this in terms of increased workload would depend in part on the number of people in the population not currently consulting for joint problems causing 'some' disruption. One study, which described need related factors (including grade I chronic pain, i.e. low disability and low intensity) stratified by knee consultation status suggests that out of the total number of patients who had grade I chronic knee pain (n=400)

82% of them (n=326) did not consult their GP (in the past 18 months) for knee pain (Bedson et al. 2007). Thus there might be a substantial proportion of patients with less disabling knee pain who could be encouraged to consult the GP for joint pain if the GP attitude was to change to 'legitimate health problem, requires treatment'.

The importance of the GP attitude is consistent with some observational and qualitative research. There is evidence to suggest that patients with high levels of hip/knee pain were reluctant to request treatment because of an attitude displayed by the GP that arthritis was part of normal ageing and that there was little that could be offered to help (Sanders et al. 2004). Although many respondents had previously consulted GPs, their experiences were negative, with GPs appearing to confirm the lack of effective treatment and rarely offering referral to secondary care (Donovan & Blake 2000, Sanders et al. 2004). There is also evidence to suggest that respondents perceive a lack of information and support for OA, especially in terms of pain management and coping with daily activities (Victor et al. 2004) and that the GP does not take OA seriously (Alami et al. 2011). Many found difficulties in communicating with doctors and some were extremely dissatisfied with the service they had received – many reported that their doctor ignored their symptoms and reinforced the view that OA was a normal part of ageing and patients felt that they were perceived as a burden upon the NHS (Gignac et al. 2006).

This finding reflects the basis of the declaration of Montreal (2010), which states that it is the *'right of people in pain to acknowledgment of their pain and to be informed about how it can be assessed and managed'.* This document emphasises that *'chronic pain with or without diagnosis is highly stigmatized'*

(International Association for the Study of Pain 2010, p.1). If patients with joint pain are vulnerable to stigmatization, the GP has the potential to act as a mediator between society and the patient in removing that stigma. Rather than continuing to perpetuate and reinforce the idea that joint pain is an inevitable part of ageing and is something to be accepted the GP has the capacity to shift that perception. Up until now it was not clear precisely how important the GP attitude might be in the decision to consult the GP for joint pain. The current study confirms the findings in qualitative research and goes further by quantifying the relative importance of the GPs attitude compared to other factors. The relative importance of the GP attitude is particularly informative compared to other service factors offered in primary care management.

The importance of GP treatments/investigations (relative importance of 14%) was moderate and similar to that of GP treatments/investigations. However, to date this has been a blind spot in traditional epidemiological studies, where studies have focused predominantly on clinical attributes (Bedson et al. 2007). This finding is consistent with qualitative research which suggests that the level of GP treatments/investigations is important to patients (Jinks et al. 2007, Sanders et al. 2004) especially early on in the course of OA (Mann & Gooberman-Hill 2011). It reflects quantitative research into knee OA which suggests that there is a focus on pharmaceutical care options rather than other treatments that patients prefer (Tallon et al. 2000).

However, it contradicts other research that suggests that a large proportion of individuals with OA symptoms do not consult healthcare professionals, despite the availability of a range of treatment options (Zhang et al. 2007). This study has

shown that people with joint pain would be more likely to present their problem to the GP if they felt they would get a thorough physical examination with the option of involvement of the wider multidisciplinary team in its management rather than a basic service of verbal advice and prescribed analgesia. This is consistent with current recommendations (National collaborating centre for chronic conditions 2008) and qualitative studies (Mann & Gooberman-Hill 2011). The addition of further investigations (x-rays, blood tests) and a promising new treatment yielded a small increment in utilities but was less dramatic than the step up from a basic service to an optimal one. This is perhaps just as well since there is general agreement that x-rays and other special investigations are not routinely needed to confirm the diagnosis of OA (Zhang et al. 2010a). An increase in consultations for OA would be expected if a promising new treatment was made available on the NHS.

The importance of primary care management (relative importance of 13%) was moderate and similar to the importance of comorbidities and GP treatments/investigations. This is consistent with research that suggests that the level of primary care management is important to patients (Sanders et al. 2004, Tallon et al. 2000). Mann and Gooberman-Hill (2011) suggest that patients with hip/knee OA require increased attention, information, and primary care management, especially in the early stages.

The relative importance of ‘competing comorbidities’ (13%) was also moderate. This contradicts some evidence that suggests that comorbid illness, when recognised as important, may result in nonconsultation for knee problems (Bedson et al. 2007). Those who identified their knee pain as a health priority were

more likely to have consulted about their knee than those who named other comorbid illness as their priority (Bedson et al. 2007). Conversely, it supports other evidence, such as Thorstensson et al. (2009) who suggest that comorbidities are important determinants of help-seeking. The PPCBC estimates suggest that other health problems may in fact increase respondents' propensity to consult for their joint pain.

However, despite extensive developmental testing with the RUG, it is unclear whether respondents fully understood that other physical health problems referred specifically to comorbidity unrelated to the joint problem (e.g. diabetes, heart problems). Respondents might have included things like the joint being hot and swollen, the knee giving way or locking in the same category of other physical health problems.

Of the attributes included in the PPCBC survey, the patient's decision to consult the GP for joint pain was least influenced by pain characteristics (relative importance of 5%). This is consistent with current research because although pain severity has long been positively associated with help-seeking in this population (Bedson et al. 2007), recent evidence suggests that severe pain is not associated with consultation (Jordan et al. 2006).

Interestingly, the study upon which the pain characteristics attribute levels were based predicted that for hip/knee OA, a higher frequency of unpredictable joint pain was significantly associated with decreasing quality of life (Hawker et al. 2010). Unpredictable knee pain was also associated with greater pain and functional disability (Hawker et al. 2008). So, it might be hypothesised that these would be quite strongly related to the decision to consult. However, the degree to

which these two pain types predict healthcare utilisation is ongoing and as yet unpublished work (by Hawker et al.). Thus there is limited empirical evidence against which to compare the present study results. This study found that the patient's decision to consult the GP for joint pain was least influenced by pain characteristics (relative importance of 5%) but this might be because this was relative to other important determinants of the decision to consult (such as pain disruption) as well as incentives of service that were on offer – determinants that were not necessarily covered in the studies of Hawker et al. (2008).

However, while evidence about the association of pain severity and help-seeking is contradictory it is unclear to what extent pain has been measured in the same way as in this (conjoint) study. Thorstensson et al. (2009) produced adjusted odds ratios for pain severity and mobility and found that disability is a greater determinant of help-seeking than pain severity for community dwelling adults with chronic pain in hip/knee. Thus it appears that these attributes have not been placed side by side in a choice-based conjoint task before now.

The attribute with the highest mean importance rating, GP attitude, was significantly different from all other attributes. Therefore, it might be interpreted that respondents understood the difference between the other attributes and the highest ranked mean (GP attitude) and were able to discriminate between them. There appears to be three distinct groups based on the means: i) the 3 service factors (GP attitude, primary care management and GP assessments) rated highest, ii) pain characteristics and pain disruption and iii) comorbidity as rated the lowest.

Evidence of floor and ceiling effects in some attributes reinforced the finding that respondents are already able to discriminate between attributes (see Figure 8.2 on page 230). There is evidence of a floor effect in the rating of attribute 3 (comorbidity) and a ceiling effect in attributes 4 (GP assessment), 5 (PCM) and 6 (GP attitude). This suggests that the direct ratings scale is able to give an average relative importance of the six attributes and far from being all rated the same, some attributes did appear to be more important to respondents. Thus in the direct rating exercise respondents were already able to discriminate between the importance of attributes (for example, the importance of comorbidity had a floor effect suggesting respondents felt strongly that this was least important). This is not consistent with conclusions drawn in some studies that direct ratings allow all attributes to be rated equally (Hjelmgren & Anell 2007).

It appears that when fully specified (in the questionnaire) direct ratings do appear to permit some degree of discrimination. However, the order of the relative importance of the attributes were different when estimated using conjoint analysis. The most prominent difference between the direct rating and CBC estimates (ranked in order) was for attribute 2, the level of disruption to everyday life. Thus, attribute two became more important in the CBC exercise than it was in the direct rating.

The striking difference between direct rating and conjoint findings (relative importance of pain disruption to everyday life) may be due to the premise that participants wouldn't consider consulting the GP if there was no disruption to their daily activities and therefore did not consider this a plausible scenario whilst doing their direct ratings. This study may be consistent with the finding by Phillips et al.

(2002a) that conjoint analysis is useful at forcing respondents to think more deeply. Overall, this study reinforces the finding that there is discordance between direct rating of attributes and CA utilities (Hundley & Ryan 2004).

Previous studies suggest that CA has advantages over direct rating scales, such as not having floor or ceiling effects (Bridges et al. 2010). However, it is unclear whether CA prevents ceiling and floor effects because if one attribute level is clearly important in CA, the scenario with that attribute level may always be selected regardless of what levels the other attributes are stated to be.

Hundley and Ryan (2004) argue that both methods are valid but that respondents rated all attributes (of intra-partum care) as the same and gave a socially acceptable response in the rating scale, whereas the discrete choice experiment elicited a more honest response. It is not clear whether this is the case in this study because it is unclear which attributes might be socially sensitive. If it were assumed that the GP attitude is socially sensitive, in this case the Hundley and Ryan (2004) finding is not consistent with this study because respondents rated the GP's attitude as the most important attribute in the direct rating, yet in the PPCBC task it appeared to be second to pain disruption.

However, a recent comparison of rating and CA methods to measure consumer preferences for hearing aid attributes suggested a high level of concordance between the two methods (Bridges et al. 2010). Bridges et al. (2010) suggest that this was because researchers constrained the rating of attributes to the individual attribute levels, whereas previous studies may have exaggerated differences between the two methods because they have typically asked respondents to rate attributes and not the individual levels which characterise the

attributes. In this study although the six attributes were specified and described in the rating exercise, they were not directly rated (and the 15 levels covering these six attribute levels were not rated directly). This might explain the discordance between the direct rating and CA estimates.

8.6.3 Strengths and weaknesses of the study

A main strength of this study was that a thorough and rigorous process of developmental work was conducted. This ensured a degree of comprehensibility and acceptability of the choice tasks and that salient attribute levels were selected.

Some theoretical expectations were also met, for example, the greater relative importance of pain disruption over pain severity on the decision to consult. The partworth utilities confirmed a priori expectations because they increased in utility as the attribute levels progressed.

This is the only study to the researcher's knowledge that combines hypothetical clinical need factors with service factors in order to look at a relative propensity to consult rather than simply what people want from primary care services. One disadvantage of the approach taken in this study is that it is difficult to compare this with other studies that have taken the illness 'scenario' approach (as defined in chapter 2 on page 32 based on an example from Caldow et al. 2007). For example, Cheraghi-Sohi et al. (2007) were interested in the degree to which the nature of the condition influenced decision-making, and decided on two key dimensions - perceived urgency, and degree to which psychological factors were implicated. They chose generic scenarios that potentially related to all respondents. Instead this study recruited patients with one condition (joint pain)

and asked them to imagine they had other minor or major physical health problems.

Clear limitations include that the PPCBC design was inefficient and likely to be underpowered in terms of estimating interactions. This meant that some of the issues raised were not able to be understood in any greater detail. For example, the issue raised of whether the difference between GP attitude matters more depending on the level of pain disruption (clinical need). In order to test interactions such as these precisely a full-profile design would be needed. However, the strong feedback in the developmental work from patients was that the presentation of six attribute levels in this study would be too burdensome. Unlike product attributes in marketing research (e.g. cost, colour, size) it is difficult to reduce concepts like the management options for OA or GP attitude to brief, universally understood and relevant descriptions.

In the final design tests of respondent consistency were not included. Therefore, it was not possible to examine the quality of the responses (for example, rationality). However, it was decided that there was enough evidence from the developmental work with the RUG and in the CA literature to suggest that preferences may be constructed at the time of elicitation rather than pre-existing – so data was not removed based on tests of consistency in this study. The data from this study could be used to investigate how choices may have changed (and were being constructed as they learnt about their preferences) during the CBC choice tasks.

Finally, predicting real-world decisions may be limited if the range of attributes that may be salient to subjects were not included in the choice

questionnaire. However, it was clear that six attributes was the absolute limit for the RUG in order to ensure an acceptable level of respondent efficiency while completing the tasks.

8.6.4 Implications for clinical practice and future research

Service factors appear to be just as influential as the most important need factor (i.e. pain disruption to everyday life) in determining GP consultation. This suggests that observational epidemiological studies that have traditionally focused only on patient/problem attributes will have obtained only a partial view of the determinants of consultation.

The clinical implication of the findings are more complex and depend to some extent on whether nonconsultation for joint pain/OA is regarded as a positive or a negative phenomenon. The emphasis (as outlined in chapter 1) in this thesis is not on a simple need to increase GP consultation. If there are people with severe pain who do not consult their GP, it does not necessarily indicate unmet need for healthcare. It may alternatively indicate successful adaptability and self-management (Huber et al. 2011) and a healthy independence from the system. However, when considering the impact of pain upon nonconsulters the emphasis is upon missed opportunities for lifestyle and self-management advice (i.e. secondary prevention) – something that can be offered in GP consultations.

This study suggests that consultation may be encouraged by improving the rigour of clinical assessment by the GP, involving the multidisciplinary primary care management team, and most of all sending out the message that ‘joint pain is a legitimate health problem that requires treatment’.

One implication of this study is that if the patient believes that the GP perceives that they have a legitimate problem they are more likely to consult. Thus, if GPs conveyed a more positive legitimising attitude about the treatment of peoples joint problems, there may be a reduction in the tendency for older adults to normalise their disease (Gignac et al. 2006, Kee 1998) – where pain and stiffness are perceived as common and expected as a natural part of ageing, and not indicative of a problem that can be treated. This may then lead to a higher degree of 'expressed need' in the GP consultation for people with less pain disruption to their daily lives (including those with only some pain disruption). Although this study does not directly address many of the issues that it raises, there is nevertheless a fairly consistent picture that while there are a range of effective interventions for OA, many are under-used (Porcheret et al. 2007, Steel et al. 2008, Peat & Thomas 2009, Li et al. 2011).

The findings in this thesis imply that patients are sensitive towards non-legitimising attitudes and need to feel that their joint pain is acknowledged and taken seriously by their GP. Perhaps a less traditional 'narrative' of healing is required (Egnew 2009) of the GP in order to work with joint pain patients, because in the scenario of joint pain the patient is experiencing a chronic illness. Chronic illness, by definition, cannot be cured, so the technology of medicine is limited to palliation and protection of function rather than cure. So, the approach required may be one of accompaniment on the chronic illness journey and encouragement as the patient struggles with the demoralisation of chronic illness and disability. This is in contrast to the more conventional narrative of 'fixing' that accompanies cure of disease (Egnew 2009). For example, in the context of joint pain, the GP

would refer to the range of treatments and self-management information available rather than perpetuating the attitude that joint pain requires acceptance only.

In contrast to cross-sectional or longitudinal observational studies, which model actual consultation behaviour in people 'as they are', PPCBC is a stated preference technique, i.e. a hypothetical behaviour. It is therefore unable to determine what proportion of people with joint pain feel their GP has a 'normal ageing-accept it' attitude or what proportion of patients are currently offered optimal care. According to a national population survey of adults aged 50 or more in England the percentage of indicated care received by eligible participants was low for osteoarthritis care at 29%. Out of the 993 times that quality indicators were eligible these were only achieved 288 times (95% CI = 26.0% to 31.9%) (Steel et al. 2008). However, the GP attitude might be more difficult to assess. It might also be that the GP attitude is already changing to become more legitimising in consultations with patients, so it is unclear how many people might be encouraged to consult based on making service provision changes (such as the GP attitude). Therefore, the clinical practice implications of this study cannot be fully understood without knowing the actual frequency of the attributes and levels in the real world. The lack of external validity in the reviewed studies in the systematic review (chapter three) reflect the wider literature, whereby studies into external validity of CA in health applications are limited (Mark & Swait 2004). However, one of the main attractions of the conjoint technique is that new and modified service configurations can be included and patients' preferences elicited on them. Real-world data may not exist on these modified service configurations.

Moreover, it is important to remember that in this design the 'normal ageing-accept it' GP attitude is rooted in the patients perception and may not accurately reflect the GP's own true attitude. It may be that the patient themselves have a 'normal ageing-accept it' attitude towards their own joint pain. Table 7.6 suggested that 17%, n=144, of the sample believed that ageing was the cause of their joint pain, so it is clearly an attitude that is shared by a significant minority of patients. Nevertheless, perhaps shifting this 'normal ageing-accept it' attitude towards joint pain is something that is potentially possible via the interpersonal communication between patient and health professional.

Future research would benefit from a deeper understanding of what proportion of people with joint pain feel their GP has a 'normal ageing-accept it' attitude and what proportion of patients are currently offered optimal care. A greater understanding of the actual frequency of the attributes and levels in the real world might give better insight into the clinical practice implications of this study.

The finding that a 'legitimating" GP attitude incentivises consultation requires deeper investigation and questioning. An important question is: is it justified for a GP to promote the attitude that joint pain is a legitimate health problem that requires treatment? Although there is a point at which this study does not address these larger questions this thesis argues that at present there is evidence of effectiveness for rather a broad range of treatments (see Figure 1.1 on page 5), many of which are under-used. This thesis emphasises that when considering the impact and severity of pain upon nonconsulters the emphasis is upon missed opportunities for lifestyle, self-management advice and secondary

prevention (as well as factors like communication and information-delivery skills) that can be offered in GP consultations. Thus, if the term ‘treatment’ is understood by patients within a broader umbrella of ‘disease management’ - as is propagated in the Model for Effective Chronic Illness Care (Wagner et al. 1999) - it is justified for the GP to promote the attitude that joint pain is a legitimate health problem that requires ‘treatment’.

The potential pitfall of promoting the attitude that joint pain is a legitimate health problem that requires treatment is that OA may become over-medicalised and treatment might be understood by patients as synonymous with ‘disease cure’. Promoting that there is a reliable ‘cure’ to patients might lead to potential harms such as collusion with the GP (i.e. encouraging them to continue to ‘chase the treatment’ to relieve pain). For example, there is evidence that struggling to get control of pain is related to (and predictive of) problematic functioning (e.g. greater disability and worse depression) (McCracken et al. 2005). The message promoted to OA patients thus needs to be to seek pain relief but not when it fails to be of benefit. When pain relief is possible, then patients should utilise what brings relief, but when continued attempts continue to fail, another course of action is required (e.g. cognitive-behavioural interventions). Continuing attempts are likely to lead to further struggling. One could consider all attempts at pain reduction as a form of struggling. If this is true, then one could consider all pursuit of treatment as a form of struggling.

This leads to the question of whether medicalising joint pain is helpful to patients and leads to more effective care. Perhaps referral is not necessary for some patients and instead the GPs goals may need to be focussed on trying to

protect/increase function. For example, patients may already have good self-management strategies but may seek clarification (for example, they may not be sure whether it is acceptable to go to the gymnasium with chronic pain or not). Similarly, the treatment might also entail the development of a therapeutic alliance with the GP with ongoing assurance that the patient will be encouraged and given access to resources during the progression of their disease. So, the narrative becomes one of support and empowerment on the chronic illness journey and not simply on cure of a disease (Egnew 2009). However, it is unclear how achievable this is within the time restraints of general practice. Future research might investiate whether GPs will want to sign up to this type of training.

8.7 Conclusion

During the course of testing PPCBC in a novel field, service factors appear to be just as influential as the most important need determinant of consultation (i.e. pain disruption) - something that has previously been a blind spot in epidemiological studies. Believing the GP would regard joint pain as 'part of the normal ageing process that one just has to accept' is a strong disincentive to seeking help, potentially outweighing other aspects of quality of care.

The next chapter will focus on exploring whether there are subgroups of respondents with different preferences for GP consultation.

9 Chapter Nine: Results III: subgroup analysis

9.1 Introduction

Chapter 8 examined the trade-offs made by patients in the decision to consult the GP for joint pain, in response to two clinically relevant questions. It was concluded that service factors appear to be just as influential as the most important need determinant of consultation (i.e. pain disruption). Believing the GP would regard joint pain as 'part of the normal ageing process that one just has to accept' is a strong disincentive to seeking help, potentially outweighing other aspects of quality of care.

In this chapter the focus is on identifying subgroups of the population with similar preferences for GP consultation. The aggregate analysis in chapter 8 provides an overall average for the respondents. However, this approach to analysis ignores possible subgroups of patient within the data which differ in their preferences on the relative importance of the attributes. This analysis will help to understand whether there are systematic differences in preferences that can be explained by other factors. In the population, for example, it is unclear whether gender (male versus female) or length of time with the joint problem influences preferences for GP consultation.

9.2 Aims

Using data gathered from the main PPCBC study in community-dwelling adults aged 50 years and over with joint pain, the aims of the present analyses were:

To identify distinct subgroups of respondents based on the choice data (preferences) and compare them with respect to descriptive characteristics.

The method of data collection and statistical analysis are presented in section 6.4.2 (see page 182).

9.3 Methods

9.3.1 Participants

A random sample of 250 out of the total of 863 respondents to the conjoint survey were included in this analysis.

9.3.2 Data

The main data were the responses to the 10 choice tasks. In addition, self-reported information from the conjoint survey questionnaire and from the previously completed NorStOP 6 year follow-up questionnaire was used to describe the characteristics of the subgroups. This included:

- demographic and socioeconomic characteristics – age, gender, primary care practice, marital status, living arrangement, employment status and occupational class.
- joint pain – hip, knee or hand pain in past 12 months, average pain intensity of joint pain in past six months (0-10 NRS), time since onset, episode duration, time since first GP consultation, healthcare professional consultations for joint pain in past 12 months (GP, hospital specialist, physiotherapist, OT, nurse, chiropractor, osteopath, acupuncturist, homeopath, aromatherapist, other), brief illness perceptions questionnaire (BIPQ: (Broadbent et al. 2006)), one item coping strategies questionnaire (Jensen et al. 2003).

- general health - a range of 22 comorbidities, anxiety and depression (Hospital Anxiety & Depression scale (HAD:(Zigmond & Snaith 1983).

9.3.3 Statistical analysis

Latent class analysis was applied to the standardised partworth utilities for each of the 250 respondents. The analysis was ran looking at a two, three, four and five class solution. There are no fixed rules for selecting the 'best' solution from latent class analysis. For each solution the following parameters were used to inform the choice of 'best' solution. These are consistent with previous studies using latent class analysis (Dunn et al. 2004):

- The Consistent Akaike Information Criterion (CAIC) model fit statistic (lower values indicate better model fit).
- The distinctiveness and meaningfulness of the subgroups (interpretability). This was judged by the distinctiveness of the attribute importance scores and evidence of subgroups being sub-divided rather than qualitatively different new subgroups emerging as the number of subgroups increased.
- The proportion of respondents confidently (probability above 0.7) allocated to a particular class.
- There are no guidelines within the Sawtooth literature regarding ideal subgroup size, but more respondents in each subgroup size was considered preferable in addition to the above criteria.

After selecting the subgroup solution, the subgroups were then described and interpreted according to the attribute utility ranges and importance scores in the choice tasks. A wide range of demographic and health characteristics were described using simple frequencies and compared across subgroups using chi-

square and Kruskall-Wallis tests for categorical and continuous variables respectively.

Latent Class Analysis was conducted in Sawtooth Software version 7 (Inc. Orem, UT). The descriptive analyses were conducted in SPSS (Inc, Chicago, IL).

9.4 Results

9.4.1 Identification of subgroups

Table 9.1: Criteria for selecting subgroup solution

Subgroup solution	(Minimum) class size	CAIC model fit	Number of respondents with Probability of membership below 0.7 (n)
2	94, 156	3012.48	21
3	114, 63, 73	3064.04	36
4	51, 42, 113, 44	3110.24	29
5	41, 98, 49, 17, 45	3166.79	39

The three-subgroup solution was selected because overall it included the most distinctive and meaningful subgroups. The three-subgroup solution, despite having a higher rate of respondents whose probability of membership was below 0.7 (n=36) than the four-subgroup solution (n=29), had a comparably lower (better) fit statistic, i.e. the lowest CAIC model fit, and was more easily interpreted and more distinctive. The four-subgroup solution had two subgroups that were similar in terms of pain disruption so these were less distinct than the three-subgroup solution. In the four subgroup solution there was evidence of subgroups being sub-divided rather than qualitatively different new subgroups emerging as the numbers of latent classes is increased.

Therefore, latent class analysis yielded a three-subgroup solution with 114 (46%) of the participants in subgroup 1, 63 (25%) in subgroup 2 and 73 (29%) in subgroup 3.

Table 9.2: Attribute utility range and attribute importance scores (%) for 3 subgroup solutions

	Subgroup		
Attribute	1 (n=114)	2 (n=63)	3 (n=73)
1. Pain characteristics	‡ 0.8 (§7%) (¶5)	0.1 (4%) (5)	0.9 (17%) (3)
2. Disruption to everyday life	3.3 (32%) (1)	0.3(15%) (3)	1.4 (25%) (2)
3. Competing comorbidity	1.6 (16%) (4)	0.8(37%) (1)	0.4 (7%) (6)
4. GP assessment/investigations	0.5 (5%) (6)	0.0(0.2%)(6)	0.9 (16%) (4)
5. Primary care management/treatment	2.2 (22%) (2)	0.8(36%) (2)	0.5 (9%) (5)
6. GP attitude	1.8 (18%) (3)	0.2 (8%) (4)	1.4 (26%) (1)
Total attribute utility range	10.2	2.2	5.5
Strength of preferences	Strong	Weak	Intermediate
Distinct drivers	Management (not diagnosis)	Comorbidity	Diagnosis (not management)
‡ Attribute utility range = maximum score - minimum score § Attribute importance scores = (Attribute utility range/ Utility range total)*100 ¶ Rank of attribute importance scores †Total attribute utility range = Attribute utility range of all 6 attributes added together			

Pattern of attribute importance scores in subgroups

In subgroup 1 (strong subgroup) the decision to consult the GP for joint pain was most strongly influenced by the 'disruption of pain to everyday life' (32%), and least influenced by GP assessments/investigations (5%). Compared to subgroup 3 (which also included disruption as an important attribute) subgroup 1 prioritised primary care management/treatment, whereas subgroup 3 prioritised GP

assessment/investigations (16%). Compared to subgroup 3, subgroup 1 appears to reflect individuals in whom management and not diagnosis is of more concern.

In subgroup 2 (weak subgroup), comorbidity (37%) was the most important attribute and least important was GP treatments/investigations (0.2 %). Competing comorbidity (as well as primary care management/treatment) was a distinguishing attribute in subgroup 2 (37%), compared to subgroups 1 (16%) and 3 (7%). It might therefore be hypothesised that this group are older, have higher depressive symptoms and more significant comorbid health problems.

Subgroup 3 (intermediate subgroup) was most strongly influenced by GP attitude (26%), and least influenced by competing comorbidities (7%). While subgroup 1 also rated pain disruption and GP attitude as important, subgroup 3 appeared to value pain characteristics and GP assessment/investigations more than management/treatment. Compared to subgroup 1, subgroup 3 reflects individuals in whom diagnosis and not management is of more concern.

Total attribute utility range in subgroups

However, there is another way of interpreting the subgroups, which is based on the difference in the total attribute utility ranges. This refers to the attribute utility range of all six attributes added together in each subgroup and is an indication of strength of preference (i.e. the higher the total utility range the stronger the strength of preference).

Subgroup 1 had the largest total attribute utility range (10.2). Overall, subgroup 1 appears to have strong preferences. For example, the lowest range (0.5) in subgroup 1 was comparable to the highest range in subgroup 2 (0.8). Subgroup 2 had the smallest total attribute utility range (2.2). Overall, subgroup 2

appears to have weak preferences. This may suggest that respondents in subgroup 2 may be effectively indifferent to the task, believe strongly in the importance of attributes not included in the study, or not to have understood it. Subgroup 3 had an intermediate total utility range (5.5) with GP attitude (1.4) and pain disruption (1.4) as joint largest attribute range and competing comorbidity as the smallest (0.4). Overall, subgroup 3 appears to have an intermediate strength of preferences.

9.4.2 Comparative description of subgroups

Subgroup membership was not associated with gender, age, practice, marital status, living situation (see Table 9.3 overleaf) or the difficulty of imagining the hypothetical scenarios in choice tasks (based on an ordinal scale of 1= not at all hard to 5= extremely hard) (see Table 9.4 on page 258).

Table 9.3: Demographic and socioeconomic characteristics of 250 randomly selected respondents, stratified by subgroups 1-3

	Subgroup 1: Strong	Subgroup 2: Weak	Subgroup 3: Intermediate	*p* value
	n (%)	n (%)	n (%)	
N	114	63	73	
Age stratum (years):				
50-64	37 (32)	12 (19)	21 (29)	.294
65-74	41 (36)	29 (46)	33 (45)	
75+	36 (32)	22 (35)	19 (26)	
Female	69 (61)	34 (54)	36 (49)	.308
Practice:				
1	36 (32)	26 (41)	19 (26)	.253
2	35 (31)	18 (29)	29 (40)	
3	23 (20)	8 (13)	8 (11)	
4	20 (18)	11 (17)	17 (23)	
Married/cohabiting	83 (73)	40 (63)	50 (68)	.525
Lives alone	25 (22)	18 (29)	18 (25)	.559
Employed	26 (23)	6 (10)	22 (30)	.025
Socioeconomic classification:				
Higher managerial	6 (5)	1 (2)	3 (4)	.014
Higher professional	5 (4)	2 (3)	0 (<1)	
Lower managerial/professional	27 (24)	5 (8)	11 (15)	
Intermediate occupations	15 (13)	5 (8)	13 (18)	
Self-employed	6 (5)	4 (6)	4 (5)	
Lower supervisory/technical	9 (8)	4 (6)	5 (7)	
Semi-routine occupations	18 (16)	14 (22)	12 (16)	
Routine occupations	22 (19)	21 (33)	20 (27)	

Subgroup membership was significantly associated with socioeconomic class (p=.014) and employment status (p=.025). Respondents were less likely to be employed if they were in subgroup 2 (employed = 10%), whereas subgroup 1 (23%) and 3 (30%) had higher rates of employment.

Table 9.4: Difficulty rating of choice tasks from 250 randomly selected respondents for subgroups 1-3

	Subgroup 1: Strong	Subgroup 2: Weak	Subgroup 3: Intermediate	*p* value
N	114	63	73	
	Median (IQR)	Median (IQR)	Median (IQR)	
† ordinal scale for difficulty of hypothetical scenarios (1-5): Median (IQR)	2 (1,3)	2 (1,3)	2 (1,3)	.578
IQR Inter-quartile range; † Difficulty was measured on a 1-5 ordinal scale where 1 = not at all hard to imagine hypothetical scenarios and 5 = extremely hard to imagine hypothetical scenarios				

There were very few apparent differences between members of the three different subgroups for regions of pain, intensity of pain, onset of pain, pain characteristics, healthcare utilisation (see Appendix 11a), self-reported comorbidities, the Hospital Anxiety Depression Scale (HADS) (see Appendix 11b), illness perceptions and coping strategies (see Appendix 11c).

Subgroup membership was significantly associated with the importance rating of pain characteristics (p<.001), the level of disruption to everyday life (p<.001), competing comorbidity (p=.039), primary care management/treatment (p=.021), GP assessments/investigations (p=.028) and GP attitude (p=.042) (see Table 9.5 overleaf).

Table 9.5: Direct rating of attributes from 250 randomly selected respondents for subgroups 1-3

	Subgroup 1: Strong		Subgroup 2: Weak		Subgroup 3: Intermediate		*p* value
N	114		63		73		
	Median (IQR)	Mean (rank)	Median (IQR)	Mean (rank)	Median (IQR)	Mean (rank)	
1. Pain characteristics	6 (4,8) (5)	6.14 (5)	4 (3,6) (6)	4.30 (4)	6 (4,8) (4)	5.87 (4)	<.001
2. Level of disruption to everyday life	8 (4,8.5) (4)	6.27 (4)	4 (2,6) (6)	4.19 (5)	6 (3,8) (4)	5.70 (5)	<.001
3. Competing comorbidity	5 (3,8) (6)	5.46 (6)	4.5 (1,6) (4)	4.07 (6)	5 (2,7) (6)	4.46 (6)	.039
4. GP assessments/investigations	8 (5,10) (1)	6.98 (3)	5 (3,9) (2)	5.63 (3)	7 (5,9) (2)	6.55 (3)	.028
5. Primary care management/treatment	8 (5,10) (1)	7.09 (2)	5 (4,9) (2)	5.72 (2)	7 (5,9) (2)	6.70 (2)	.021
6. GP attitude	8 (5,10) (1)	7.22 (1)	7 (4,9) (1)	6.17 (1)	8 (5,9) (1)	7.00 (1)	.042

Table 9.5 (on the previous page) shows that subgroup 2 had a slight discrepancy between means and medians in direct rating rankings. However, the problem with using the median is that this inevitably leads to equal rankings. It was thus decided that the mean was more appropriate for ranking and comparing subgroups. Table 9.5 shows that in terms of the direct rating rankings subgroups 2 and 3 look to be consistent with the whole sample (as in Table 8.10 on page 230) (i.e. the GP / primary care attributes are most important and comorbidity is least important). Subgroup 1 is very similar to subgroups 2 and 3 except that pain disruption is more important than pain characteristics in subgroup 1.

However, the mean scores (for direct rating of attribute) were systematically higher in subgroup 1 (strong preferences) than subgroup 2 and 3. Subgroup 2 (weaker preferences) had systematically lower mean scores than the other subgroups and subgroup 3 (intermediate preferences) had systematically higher mean scores than subgroup 2 and lower mean scores than subgroup 1.

Table 9.6: Direct rating of attributes (ranked by mean scores) and conjoint utilities (ranked by relative importance scores) for subgroup 1

	Subgroup 1: Strong (n=114)		
Attributes	Direct rating rank (Mean)	Choice tasks rank (relative importance)	Increase of importance in PPCBC: e.g. +/- (in rank)
1. Pain characteristics	5 (6.14)	5 (7%)	Same (0)
2. Level of disruption to everyday life	4 (6.27)	1 (32%)	+ (3)
3. Competing comorbidity	6 (5.46)	4 (16%)	+ (2)
4. GP assessments/investigations	3 (6.98)	6 (5%)	- (3)
5. Primary care management/treatment	2 (7.09)	2 (22%)	Same (0)
6. GP attitude	1 (7.22)	3 (18%)	- (2)

Table 9.6 shows that in subgroup 1 (strong preferences) the highest rating (ranked by mean importance score) attribute (GP attitude) and the third highest (GP assessment) became less important in the CA exercise (i.e. they were traded off). Two attributes (pain disruption and comorbidity) increased in importance and two attributes (pain characteristics and primary care management) remained the same. It is clear that respondents were forced to trade against the highest ranking attribute from the rating results when completing the CA exercise. Interestingly, the second highest ranking from the direct ratings (primary care management) remained a firm priority. It appears that pain disruption became more important in the CA exercise: a finding that is similar to that found in the whole sample.

Table 9.7: Direct rating of attributes (ranked by mean scores) and conjoint utilities (ranked by relative importance scores) for subgroup 2

	Subgroup 2: Weak (n=63)		
Attributes	Direct rating rank (Mean)	Choice tasks rank (relative importance)	Increase of importance in PPCBC: e.g. +/- (in rank)
1. Pain characteristics	4 (4.30)	5 (4%)	- (1)
2. Level of disruption to everyday life	5 (4.19)	3 (15%)	+ (2)
3. Competing comorbidity	6 (4.07)	1 (37%)	+ (5)
4. GP assessments/investigations	3 (5.63)	6(0.2%)	- (3)
5. Primary care management/treatment	2 (5.72)	2 (36%)	Same (0)
6. GP attitude	1 (6.17)	4 (8%)	- (3)

Table 9.7 shows that in subgroup 2 (weak preferences) the highest ranked attribute from the rating results (GP Attitude) and pain characteristics became less important and were traded off in the CA exercise. Two attributes increased in importance (pain disruption and comorbidity) with comorbidity increasing dramatically and one attribute remained the same (GP assessment).

Table 9.8: Direct rating of attributes (ranked by mean scores) and conjoint utilities (ranked by relative importance scores) stratified by subgroup 3

	Subgroup 3: Intermediate (n=73)		
Attributes	Direct rating rank (Mean)	Choice tasks rank (relative importance)	Increase of importance in PPCBC: e.g. +/- (in rank)
1. Pain characteristics	4 (5.87)	3 (17%)	+ (1)
2. Level of disruption to everyday life	5 (5.70)	2 (25%)	+ (3)
3. Competing comorbidity	6 (4.46)	6 (7%)	Same (0)
4. GP assessments/investigations	3 (6.55)	4 (16%)	- (1)
5. Primary care management/treatment	2 (6.70)	5 (9%)	- (3)
6. GP attitude	1 (7.00)	1 (26%)	Same (0)

Table 9.8 shows that in subgroup 3 (intermediate preferences) the highest and the lowest ranked attributes from the rating results (GP attitude and comorbidity) remained the same. GP attitude firmly remained the priority for subgroup 3. Two attributes became less important and were traded off in the CA exercise (GP Assessment and primary care management) and two became more important during the CA exercise (pain characteristics and level of disruption).

9.5 Discussion

9.5.1 Principal findings

There appeared to be no distinct different 'types' of person based on their different patterns of attribute importance within the three subgroups, although there was an ability to identify comorbidity, diagnosis and management driven subgroups of respondents. Instead the strength of preference appeared to be a distinct axis of the three subgroups identified. Thus the strength of preference appeared to be generally independent of all descriptive characteristics (apart from major differences between the subgroups in terms of their direct rating of the six attributes).

Respondents in subgroup 1 (strong preferences) had systematically stronger direct importance ratings for the attributes in the decision to consult the GP for joint pain. Respondents in subgroup 2 had systematically weaker direct importance ratings for the attributes. Respondents in subgroup 3 systematically had intermediate direct importance ratings for the attributes in the decision to consult the GP for joint pain.

These results reinforce the findings in section 8.6.2 that when fully specified direct ratings do appear to permit some degree of discrimination. However, the order of the relative importance of the attributes for the subgroups were different when estimated using conjoint analysis. Thus from the direct ratings there is no sense of the different types of preferences but from the conjoint exercise it emerges that there are some qualitative differences between the three subgroups (i.e. the management-driven (subgroup 1), the comorbidity-driven (subgroup 2) and diagnosis-driven (subgroup 3) subgroups).

The results suggest that CA might be useful at forcing people with weaker preferences (as in subgroup 2) to trade, because initially respondents rated attributes low but then during the CA exercise some attributes became more important (in this case comorbidity became most important). Only one attribute remained the same. There is also evidence that CA is less useful at forcing people with positive direct ratings (as in strong subgroup 1 and 3). In strong subgroup 1 the highest and lowest rated attributes were traded-off and became less/more important, although the second highest and second lowest appeared to remain a firm priority. In the intermediate subgroup 3 the highest and lowest rated attributes

remained the same. Particularly, GP attitude remained the firm priority and comorbidity the lowest priority in subgroup 3.

9.5.2 Interpreting the principal findings in the context of previous studies

Previous studies have traditionally used subgroup analysis to look at distinct types of person based on their different patterns of attribute importance (Cunningham et al. 2008). Few have looked directly at the strength of preferences in this way. However, there has been some research (not based on subgroup analysis) to suggest that CA exercises elicit a more honest (less socially desirable response) (Hundley & Ryan 2004, Hjelmgren & Anell 2007).

Bridges et al. (2010) suggest that previous studies comparing the rating and conjoint methods may have exaggerated differences between them because the individual attribute levels were not specified. Only the attributes as a whole were individually interpreted. In the present study using subgroup analysis there was clear discordance between the rating and CA and this may be as a result of not asking respondents to rate the individual attribute levels (Bridges et al. 2010). Future research should look into the discrepancies between these different approaches. Although the attribute levels were described in the rating exercise to explain what the attributes referred to, perhaps if they were to rate each level individually results might be more concordant.

The strong subgroup (1) suggests that there are a small portion of people who are management driven in terms of consultation. Subgroup one's preferences are consistent with research done which suggest that patients with hip or knee OA, especially in the early stages require greater attention, better information, and

more proactive primary care management (Sanders et al. 2004, Mann & Gooberman-Hill 2011). Identifying a management driven subgroup may contribute to an understanding of why respondents are not consulting, despite having severe pain (clinical need). Perhaps these patients do not expect to receive much support by way of management. However, this is not something that can be proven within the remit of this thesis.

The weak subgroup (2) preferences suggest that there are a small proportion of people for whom comorbidities appear to encourage GP consultation. The preferences of this subgroup are consistent with research that suggests that a positive association between number of comorbidities and consultation with a GP and other healthcare professionals was found (Thorstensson et al. 2009). It contradicts evidence which suggests that help-seeking for comorbidities does not increase consultation for knee pain (Bedson et al. 2007), This subgroup supports this research and argues against OA being simply pushed off the agenda, by patients at least. However, although this subgroup was 'comorbidity-driven' there was no significant difference in terms of actual self-reported comorbidity nor duration of the problem (i.e. it might be expected that the duration of the problem was lower in this group). There was no significant difference in terms of proportion of people with comorbidities between the groups, although 32% of respondents in subgroup 2 had eyesight problems compared with 21 and 22% respectively for the other two groups. It might be the case that this subgroup were indifferent to the attributes used in the task and would not likely be driven to consult for clinical need or service attributes related specifically to the joint pain itself. Thus other health

problems became the attribute most influential in their decision to consult for their joint pain, though they might not currently have comorbidites.

The intermediate subgroup (3) suggests that there are a small proportion of people whose decision to consult may be strongly influenced by concerns about diagnosis. This subgroup's preferences are consistent with the qualitative developmental research done in this thesis (see chapter 4 on page 129) that suggests that determinants of consultation include the promise of an X-ray and diagnosis from a rheumatologist. However, although this subgroup was diagnosis-driven there was no difference in terms of actual self-reported consultation for joint pain in past 12 months with a range of health professionals.

9.5.3 Strengths and weaknesses of the study

The subgroup analysis failed to identify subgroups that were distinct in patient characteristics. This might be linked to the sample frame, which was drawn from participants in a recent local population cohort with joint pain and was selected for convenience and to maximise response. Thus, the groups most likely to be under-represented include older adults. In this way the lack of an observed association in the data (e.g. an association between GP attitude and age) may be due to the limited generalisability of the sample.

In general statistical literature it is suggested that researchers use as few significance tests as possible in order to reduce the risk of type I error (i.e. the risk of obtaining chance significance is increased by multiple testing) (Bowling, 2005). In this analysis the issue of multiple testing is apparent. They involve numerous variables being cross-tabulated in search of possible associations. Thus, even if none of the descriptive characteristics tested are associated with the subgroup

membership, chance will determine that roughly in 50 out of 1000 cross tabulations a significant association will be found at the ≤0.05 level (Bowling, 2005). Therefore, the findings from these analyses were treated modestly and in the light of the exploratory nature of the analyses.

Moreover, the subgroup analysis was limited to a sample of 250 respondents due to restrictions in the software license. There is a possibility that this was insufficient to produce stable subgroups, or detect differences in the descriptive characteristics. One of the ways of addressing this would be to repeatedly draw random samples of 250 and rerun the latent class analysis to see if a consistent solution is repeatedly drawn from the 250 persons from the 863 respondents. Alternative strategies to latent class analysis for detecting subgroups could also be explored.

One strength of subgroup analysis was that it detected strong, weak and intermediate preference groups as based on the total utility ranges. It also identified the strength of preferences of the subgroups and these could be used to compare the direct rating scale estimates, shedding some light on the discordance between the two different methods.

9.5.4 Implications for clinical practice and future research

In general the findings do not provide strong evidence in favour of distinct types of preferences profile. However, this study does seem to identify comorbidity, diagnosis and management driven subgroups of respondents. However, it did not allow enough individuals with these different preferences to be identified. With more rigorous testing to find certain types of people within these subgroups this information could help the NHS design materials consistent with

each subgroup's preferences, identify the components of individual management plans and estimate demand for different resources. The management driven section, for example, implies that there is a need for greater awareness about the availability of optimal services (i.e. attribute level 2 in GP assessment and primary care management attributes) and ensuring patients are aware of them.

As an approach to understanding how different types of patients decide to initiate a GP consultation, the use of this method requires more investigation. In the present study it was clear that the strength of preference was a major factor behind respondent heterogeneity. Strength of preference reflects the degree of error (i.e. greater error leads to a weaker strength of preference). Greater error can be due to aspects such as greater respondent fatigue, respondents not understanding the task or being inconsistent in their answers. However, Table 9.4 shows that although subgroups had varying strengths of preference, the median difficulty rating for completing the hypothetical choice tasks did not differ across subgroups. Thus the median number for subgroup 2 (the weaker strength of preference) did not differ from the other two subgroups in terms of difficulty to complete the choice tasks. Future research might compare PPCBC (designed to only present a subset of attributes at one time) with adaptive choice-based conjoint (a computer-based adaptive design enabling more complex attributes to be handled without over-burdening respondents). This might be an opportunity to see how strength of preference might increase or decrease according to different attribute levels and CA designs.

Subgroup analysis can clearly be used to compare CA utilities with other preference elicitation methods. Future research into the concordance between

direct rating, conjoint and direct ranking scales would perhaps be useful. In this study the ratings were ranked (by mean scores) and although this shows the benefits of CA (i.e. respondents are forced to trade) it would be interesting to see if CA have much greater benefits than direct ranking approaches, since in a direct ranking exercise respondents are also forced to prioritise. A comparison of CA with ranking scales in this context might enable a greater understanding of the process of trading-off and prioritising in the face of a combination of different combinations of choices.

This PPCBC questionnaire reflects respondent preferences at a specific point in the NHS delivery process and utilities may shift if the act of consultation exposes patients to new information or increased outcome expectations. The test retest reliability of PPCBC at different time points before and after consultation should be an area for future research.

There is much research which suggests that Hierarchical Bayes analysis provides more accurate and stable estimates of utilities for the purposes of subgroup analysis. This could be addressed in future research (Orme 2006).

9.5.5 Conclusion

An exploratory approach to identifying the characteristics of patients with different preferences for GP consultation was taken. Results suggest that there are no distinct different ‘types’ of person based on their different patterns of attribute importance.

Instead, strong, weak and intermediate preference groups were identified based on strength of preference. When compared with the direct rating scores the strong, weak and intermediate subgroups identified corresponded with the mean

importance rating scores. However, the order of the relative importance of attributes was different in the CA exercise. CA might be useful at forcing people who have weaker strengths of preference to trade attributes, but is less useful in respondents who have stronger preferences.

The next chapter reviews the progress towards achieving the aims of the thesis, outlined in section 2.5 in chapter 2 (see page 51). This focuses on the development of the PPCBC questionnaire within the context of the determinants of GP consultation, and its potential usefulness in future research in this field.

10 Chapter Ten: Discussion, conclusions and recommendations

10.1 Introduction

This thesis arose from a series of findings from observational epidemiological studies and qualitative studies and an interest in understanding patient decision-making processes. Previous studies had found that a substantial minority of older adults with joint pain or symptomatic osteoarthritis did not consult general practice – the main point of contact for formal healthcare in the UK - for their problem over prolonged periods of time. Low expectations (Stoller et al. 1993, Sanders et al. 2004) and sub-optimal care had also been repeatedly documented despite the publication of NICE guidelines highlighting a range of effective treatment options for osteoarthritis (National collaborating centre for chronic conditions 2008) (see Figure 1.1 on page 5). It seemed likely that an understanding of the determinants of consultation would have to involve looking not just at patient characteristics and need-related factors but also at what services were on offer and how patients rated these. This challenge would require a different methodological approach to those used previously to investigate why, and which patients consult. While conventional observational epidemiological studies are able to quantify the relative importance of certain determinants, they are necessarily constrained to observed consultation and therefore to services and need-related factors “as they are”. No evaluation of hypothetical scenarios and services is possible. It has been argued in this thesis that these considerations are important for a condition like painful osteoarthritis that varies over time in its severity and impact on individuals and also for trying to evaluate preferences for different types of services, including models of care that do not yet exist. Qualitative studies undoubtedly provide more

depth of understanding into individuals' decisions to consult than epidemiological studies but a quantitative comparison of the relative importance of determinants is lacking. Conjoint analysis provided an interesting alternative approach that in principle could address some of the limitations noted in conventional techniques. However, developing and conducting a new conjoint analysis study in this field presented many challenges. These included that:

- Conjoint analysis had been little used in studies of osteoarthritis at the time of embarking on this thesis (Fraenkel et al. 2004a, Fraenkel et al. 2004b, Ratcliffe et al. 2004, Fraenkel & Fried 2008) requiring searches of literature from other health-related and non-health fields.
- There was little direct previous experience of conjoint analysis within the supervisory team, which required an ability to rapidly acquire competence in technical aspects of these methods, including the use of new bespoke software.
- There is not one but several forms of conjoint analysis and the relative merits of these different forms and modes of administration for the research questions in this thesis were unclear.
- There was considerable uncertainty about its feasibility, acceptability and ability to produce interpretable, meaningful insights into the determinants of consultation in this target population.

This chapter presents a discussion of the work undertaken in this thesis and the main findings (see Table 10.2 on page 281) followed by a critical reflection on the key decisions made and the contribution of this work to knowledge in this field.

The chapter ends with a consideration of the implications of this thesis for future research and clinical practice.

10.2 Principal findings

10.2.1 Determinants of GP consultation for joint pain in older adults

A review of the literature highlighted that there are many potential determinants of GP consultation for joint pain and the Andersen-Newman model (1995) was the most commonly used framework to organise these. Observational studies highlighted a range of factors including previous experience of healthcare, age, pain severity, pain chronicity and illness perceptions. However, while clinical need factors were important, they appear insufficient determinants of the decision to consult a healthcare professional (Jordan et al. 2006). Instead many people with longstanding joint pain become resigned to it, attributing joint problems to 'normal ageing' and having low expectations and negative experiences of primary care (Jinks et al. 2007, Sanders et al. 2004). The Andersen-Newman model (1995) - the framework extensively used to understand healthcare utilisation - has been criticised for not including aspects of technical and interpersonal skills of the healthcare provider (Bradley et al. 2002).

Conjoint analysis offered the opportunity of embedding the Andersen-Newman model (1995) in a hypothetical decision-making framework – one which could look at service configurations not currently available (including aspects of technical and interpersonal care). It was thus decided to conduct a systematic review into the applications of conjoint analysis in patient-centred care in the GP consultation in order to identify how published studies defined attributes that characterise patient-centred aspects of GP consultation (something that appeared

important in GP consultation for joint pain but was under-researched). It was clear that no previous studies had placed clinical need attributes (like pain severity) in conjunction with enabling (service) factors.

The fifteen studies included in the systematic review highlighted the importance of conducting developmental studies to inform the design of a new conjoint analysis study. Specifically these developmental studies would inform the selection of attributes, the number and wording of these attributes and their respective levels that respondents could be expected to manage, as well as the mode of administration. This thesis therefore undertook Patient Public Involvement (PPI) work with the Research Users' Group focus groups and cognitive interviews in order to maximise respondent efficiency in the final design of the Partial-Profile Choice-Based Conjoint questionnaire (as recommended by Ryan et al. 2001). This work resulted in setting several design features: 6 attributes (two 2-level attributes, four 3-level attributes), the use of partial (based on 3 attributes) rather than full-profiles, the presentation of only two profiles per choice task, and an upper limit of 10 choice tasks feasible for completion. The Advanced Design Module (see chapter 5) was then used to ensure the statistical efficiency of the design. It was decided that the main design would include: 10 random choice tasks (no holdouts), 10 questionnaire versions and estimation of main effects only.

Multinomial logit regression (based on 863 respondents, 74% response rate, 55% female, mean age 70 years, range: 58-93 years) revealed that the most important determinant of the patient's decision to consult the GP for joint pain was the extent to which pain disrupted everyday life (1.10 logits). Believing the GP would regard joint pain as 'part of the normal ageing process that one just has to

accept' is a strong disincentive to seeking help, potentially outweighing other aspects of quality of care. Their decision was less influenced by other health problems and episodes of more severe, unpredictable pain. Therefore, service factors appear to be just as influential as the most important need determinant of consultation (i.e. pain disruption). Latent class analysis failed to clearly identify distinctive subgroups of patients with different preferences. Instead, what appeared to emerge were groups who differed on their expressed strength of preference.

When nonconsultation for joint pain/OA is regarded as a negative phenomenon (i.e. indicating an under-use of potential treatments), these findings suggest that consultation might be encouraged by improving the rigour of GP assessment, involving the multidisciplinary primary care team and, most importantly, ensuring that patients feel that their joint pain is a legitimate health problem that requires treatment.

However, there are several important caveats to this conclusion. Firstly, it is based on the premise that encouraging consultation is desirable. Secondly, since partial-profile conjoint requires much larger sample sizes to accurately estimate interaction terms, it was not possible to know if consultation would be selectively encouraged in those with the greatest need or whether it might also trigger more consultations among those people who are in less need of help. Thirdly, without knowing the actual state of service provision in the UK and the perceived and actual attitudes of GPs the size of the gap between current services and optimal services cannot be known. Conjoint analysis does not provide the actual occurrence rates of the different attributes.

Overall, this study went some way towards addressing criticisms of the Andersen-Newman model (1995). Firstly, the second most important attribute to the decision to consult the GP for joint pain was the perceived GP attitude and this was something frequently omitted from the Andersen-Newman framework in the context of older adults. Moreover, conjoint analysis quantified the relative importance of clinical need and service factors simultaneously. Unfortunately, some of the predisposing and need characteristics were not drawn out in the subgroup analysis (i.e. there were no distinct classes of person within the sample who have different types of preferences) though this might be due to a limited sample of respondents being included in the analysis (n=250).

10.2.2 The design, conduct, and interpretation of conjoint analysis in this target population

By comparing the present study findings with the findings from the systematic review in chapter 3 (n=15 studies) and a broader, recent systematic review of all conjoint analysis studies (Marshall et al. 2010; n=79 included studies) it is possible to outline the present study's similarities and differences to those that had preceded it. In both systematic reviews most studies used a discrete choice experiment format, used between 3 and 16 attributes (Marshall et al. 2010) (most commonly 5 or 6), and presented between 7 and 32 scenarios to each respondent (Chapter 3). Although the number of attributes used in this study (6) was as per majority of previous studies it sought to present more information on attribute levels than usual and hence presented only a subset of attributes at one time, using a Partial-Profile Choice-Based Conjoint (see Table 10.1 on page overleaf).

Table 10.1: Comparison of main study features with previous systematic review findings

	Systematic review (Chapter 3)	Systematic review (Marshall et al. 2010)	Main study in thesis
Typical features*			
Design	DCE Full profile	DCE Full profile	CBC Partial - profile
Number of attributes	5 or 6	6	6
Number of scenarios	7-32	7-15	10
Sample size	§832	§259	863
Response rate	†60.5	Not reported	74%
† Median response rate § Mean sample size			

The present study presented 10 choice sets, which was the maximum number deemed manageable by the Research User Group. The mean sample size of the systematic review in chapter 3 and the Marshall et al. 2010 were 832 and 259 respectively. Therefore, the present study is one of the largest conjoint studies to have been conducted in the health field in recent years. The response rates across the two systematic reviews ranged from 18% to 94% (with a median response rate of 60.5% in the systematic review in chapter 3). Comparably, the response rate (74%) in the present study was good (albeit in a selected sample).

It is clear that the conclusions from the systematic review were largely consistent with Marshall (2010) – the most recently published systematic review of conjoint applications in health. However, the present study went further than the developers of the ISPOR checklist in using this as a tool for judging the quality of study and reporting. For example, the review highlighted that only one study used cognitive interviews in the development of their design (Cheraghi-Sohi et al. 2008). Therefore, cognitive interviewing was used to test the acceptability and

comprehensibility of the Partial-Profile Choice-Based Conjoint task to Research User Group members in the target population before deciding on the final design. This is discussed in greater detail in section 10.3.1 (on page 282).

Evidence that five attributes in the direct rating were statistically different from the highest ranked attribute and floor and ceiling effects in some attributes suggested that, when fully specified, direct ratings do appear to permit some degree of discrimination. However, the order of the relative importance of attributes was different based on conjoint partworths. In the conjoint exercise the extent to which pain disrupted everyday life became more important than it was in the direct rating. This may be due to the fact that participants would not consider consulting the GP if there was no disruption to their daily activities and perhaps they did not consider this a plausible scenario whilst doing their direct ratings. Arguably this is a strength of conjoint, because (as stated in chapter 1 the advantage of CA over other techniques is that) CA is a 'decompositional' method, wherein respondents evaluate scenarios composed of a combination of attribute levels, with each level explicitly stated. Thus, utilities can be estimated for each attribute level and for all possible combinations of levels, placing the levels in a context of a hypothetical scenario for respondents to consider. This to some degree goes beyond the average relative importance that direct ratings can estimate (i.e. where individual attributes are evaluated as a whole) (Phillips et al. 2002a).

The subgroup analysis did not reveal different types of people with different preferences. Instead the strength of preference was the most distinct axis of the three subgroups identified. This might be linked to the sample frame, which was

drawn from participants in a recent local population cohort with joint pain and was selected for convenience and to maximise response. Subgroup analysis suggests that conjoint analysis was able to learn about the preferences of respondents with weaker preferences, as well as (to a lesser) degree forcing respondents who had stronger preferences in the direct rating exercise to trade.

The strength of preference to some degree reflects the degree of 'error' within the choice tasks. This error could be a result of respondent fatigue, misunderstanding of the choice tasks (poor engagement) and respondent inconsistency. In this study, subgroups with a weak, intermediate and strong strength of preference emerged and this might have implications for the degree of fatigue, engagement and inconsistency experienced by respondents. It might be the case that those in the weak subgroup suffered greater respondent fatigue, greater misunderstanding of the task (poor engagement) and respondent inconsistency. However, the median score for the difficulty of completing the choice tasks did not differ across the three subgroups.

10.3 Key decisions and their implications for the interpretation of this thesis

10.3.1 Developmental studies with the Research Users' group (RUG)

One of the key decisions was to involve Research User Group members closely in the main study design. This is something that is increasingly required in applied health research (Ives et al. 2012). The findings in this thesis are consistent with much research which suggests that PPI has a significant impact upon the initial stages of research (e.g. assisting in questionnaire design and improving the sensitivity of research language) and improving response rates (Brett et al. 2010).

Table 10.2: Key findings from the thesis

Findings relevant to determinants of GP consultation for joint pain in older adults
• The extent to which pain disrupted everyday life is the most important determinant of the patient's decision to consult the GP for joint pain.
• Service factors are as influential as the most important need determinant of consultation (i.e. pain disruption). By improving the rigour of GP assessment, involving the multidisciplinary primary care team and sending out the message that joint pain is a legitimate health problem that requires treatment, consultation might be encouraged.
• Believing the GP would regard joint pain as 'part of the normal ageing process that one just has to accept' is a strong disincentive to seeking help, potentially outweighing other aspects of quality of care. Thus when included in the Andersen-Newman model of healthcare utilisation (1995), 'enabling' factors (such as attitudes towards healthcare provider) emerge as important in GP consultation.
• The predisposing and need characteristics were not drawn out in the subgroup analysis.
Methodological findings: relevant to conjoint analysis
• Partial-profile conjoint methods are well-completed in a majority of community-dwelling older adults with joint pain. Comparable to other studies, the response rate (74%) in the present study was good (albeit in a selected sample). The main limitation was an inability to estimate precise interaction effects.
• The present study is one of the largest conjoint analysis studies to have been conducted in the health field in recent years.
• The subgroup analysis failed to identify subgroups that were distinct in almost any patient characteristic. The strength of preference was the most distinct axis of the three subgroups identified.
• When attributes are fully specified, direct rating scales do permit some degree of discrimination, but they do not produce the same order of relative importance as conjoint analysis.
• Conjoint analysis might be useful at forcing people who have weaker strengths of preference to trade attributes, but is less useful in respondents who have stronger preferences.

The feedback received from the Research User Group (based on cognitive interviews and discussion groups) was highly influential for the interpretation of this thesis. Most importantly, after the negative experience of one person refusing to participate during developmental study 1, it became a priority for the remaining developmental studies to reduce and simplify the choice tasks and information presented to respondents as much as possible. This early negative experience led to the issue of respondent burden becoming a major focal point. In an attempt to reduce respondent burden three major decisions about the design were made, including the selection of attributes and levels, the selection of a partial-profile design and the selection of a postal design.

During the developmental studies with the Research User Group (in chapter 4) it was clear that there was a need to reduce the amount of information presented to respondents in order to avoid overburdening them. Van Til et al. (2009) found that extensively informed subjects (about the task and attributes) demonstrated a better understanding of the decision task (van Til et al. 2009). However, in the present study the positive effects of extensive information were weighed up against the cognitive burden placed on respondents and it was clear that in order to obtain a satisfactory response rate information needed to be reduced to a minimum.

Based on the developmental stages with the RUG it was estimated that the questionnaire would take 60 minutes to complete. However, due to the selection of a paper-based questionnaire respondents were able to return to the questionnaire and fill it out in their own time. It is not known whether respondents completed the questionnaire in one sitting or several, nor whether pain levels (or preferences) of

respondents might have changed over the course of questionnaire completion although the randomised order of the choice tasks would be expected to prevent any major order effect (if present).

Overall, the extensive testing with the Research User Group led to a 74% response rate, and a median difficulty score for the questionnaire (based on a rating scale of 1= not at all hard to 5= extremely hard) in the sample of 2 (1, 3). This is interpreted as evidence that Partial-Profile Choice-Based Conjoint is feasible in this population.

10.3.2 The selection of salient attributes and levels

A key decision in the course of this thesis was to restrict the number of salient attributes that characterise the decision to consult the GP for joint pain to six. Whilst more attributes have been used in the context of patient-centred care (Cunningham et al. 2010) the maximum in previous studies into GP consultation was 7 (Haas 2005), though it is not clear how respondents managed these. The Research User Group advised that three attributes in each scenario (a total of six pieces of information) was the maximum number they could manage (see chapter 4), a finding consistent with much psychological literature (Miller 1956). Although this approach limited the number of attributes available for inclusion in the questionnaire, it meant that patients would more likely be able to hold all of the information in their memory to ensure optimal respondent efficiency. The biggest challenge was in narrowing down the salient attributes and levels to a manageable amount for respondents and ensuring that the salient attributes were described succinctly but accurately and without compromising comprehensibility to respondents.

The content of the attributes were carefully chosen to be plausible to combine in any hypothetical scenario. Given the range of determinants of GP consultation identified in the literature search and elicited from the Research User Group it was important to select six salient attributes. However, during an oral presentation of the study findings the plausibility of the GP having a positive legitimising attitude but not offering an optimum package of management was questioned, i.e. is this a plausible combination? During the developmental stages the Research User Group patients felt it was plausible that a GP may have a positive legitimising attitude but not offer an optimal package of management and vice versa. After the study was complete the main findings presented to the Research User Group in order to demonstrate the usefulness of their input on this project and also to discuss any unexpected results with them. There were no objections from the Research User Group with the plausibility of combining the GP attitude with primary care management and GP assessments/investigations. The data from the feedback section in the questionnaire in terms of missing attributes and levels or difficulty of those included was explored and the implausibility of combining GP attitude and primary care managements GP assessment/investigations was not directly noted by any respondents (see Table 6.1 on page 177).

It was noted that although there may not be the capacity in health services to offer the optimal (and beyond optimal) service configurations described in this study, these were considered plausible to include by the Research User Group and were not disputed in the feedback sections of the questionnaire.

Three clinical need and three service factors were selected on the basis that this would give an overall balance and that neither qualitative nor epidemiological studies had previously quantified these factors simultaneously.

It may be that during the course of the questionnaire, where respondents are making specific judgements and trade-offs about what would encourage them to consult a general practitioner for joint pain, they might start to change their minds about what primary care treatment they would like to receive. However, the attributes and levels in the final scenarios were chosen to reflect the decision to consult the GP. It is likely that other attributes and levels would be relevant to the decision to consult a different healthcare professional, e.g., a physiotherapist.

10.3.3 The selection of a postal, pen-and-paper based partial-profile design

CBC was chosen because it imitates real-world choice behaviour – making it a realistic and familiar task to respondents and reducing judgment errors (Orme 2006). Partial-Profile Choice-Based Conjoint was chosen in order to increase the number of attributes that can be dealt with effectively (Orme 2006, Cunningham et al. 2008). After extensive developmental studies with the Research User Group a postal Partial-Profile Choice-Based Conjoint design was selected. Adaptive Choice-Based Conjoint (ACBC) was considered (as it has a partial-profile aspect to it) (see section 4.5) but in order for it to adapt to respondents' previous responses it must be administered by computer. The Research User Group suggested that this may be a problem for older age groups as computer access and computer literacy may be limited. In contrast to the computer design a postal format was a familiar task to this age group: a questionnaire that can be returned to and completed in their own time.

Moreover, a postal format offered access to a potentially wider sample frame, ensuring that an efficient sample size might be collected. On the other hand although respondent efficiency was enhanced and a good response rate was received, this might be at the expense of an in depth understanding of how the participants completed the choice tasks (beyond the developmental studies undertaken with the Research User Group). Overall, the aim was to ensure that respondent efficiency was optimised. This level of acceptability from the Research User Group (74% response rate) seems to suggest that Partial-Profile Choice-Based Conjoint is feasible in this population.

Overall, Partial-Profile Choice-Based Conjoint offered all of the important aspects in terms of optimum respondent efficiency. The main disadvantage of Partial-Profile Choice-Based Conjoint is that it is likely to be underpowered in terms of estimating interactions. These unobserved interactions between attributes may have serious consequences for parameter estimates. This remains a limitation of the present study which could only be addressed by a much larger Partial-Profile Choice-Based Conjoint study or a more demanding full profile study. However, it is important to note that of the fifteen reviewed studies in chapter 3 only four attempted to estimate interaction effects (Vick & Scott 1998, Scott & Vick 1999, Longo et al. 2006, Gerard et al. 2008).

10.3.4 The use of NorStOP cohorts as a sample frame for the study

Data from the North Staffordshire Osteoarthritis Project (NorStOP) were used in this thesis to identify the sample frame and recruit patients for the Partial-Profile Choice-Based Conjoint questionnaire. This also enabled previously collected data on those participants to be used in this study. The NorStOP data set

was selected for practical reasons, i.e. the NorStOP is a large, population- based, longitudinal study for which the data (including some health assessment data) were readily available and would access respondents who were 'well-disposed to' and 'willing to participate' in returning questionnaires to the ARUKPCC.

This sample frame was ideal for the purpose of this PhD project, as it meant that respondents were a sample 'well-disposed to' research. However, because there was some degree of selection involved in using conjoint analysis as a novel method, there remain questions over the generalisability of results to all older adults with joint pain (see section 7.6.3).

10.3.5 Using conjoint analysis (compared to other methods) to understand the internal process of deciding to consult the GP for joint pain

The partworth utilities derived from the choice questionnaire in this thesis present key advantages over the use of qualitative and epidemiological approaches. Decision-making is an individual dynamic process and not a fixed state. Whilst epidemiological studies typically look at people who did and did not 'consult' and test for associations between clinical need factors and qualitative studies typically look at the motivations, beliefs, and experiences of patients neither of these methods alone has been able to quantify the relative importance of determinants simultaneously. Alone, these methods cannot provide a full picture of the internal process of decision-making. Conjoint enables the relative importance of clinical and service attributes and the quantitative trade-offs to be better understood, thus complementing existing methods.

However, when modeling based on a single (average) set of utilities using aggregate multinomial logit some respondent heterogeneity is lost. Aggregated

analysis is not necessarily able to capture the different strengths of preference within the sample, whereas when looking on an individual level (at respondent heterogeneity) it is possible to observe these differences.

Latent class analysis represents one approach to understanding respondent heterogeneity and this was only partially successful in this study. Alternative approaches include Hierarchical Bayes which is reported by software developers as increasing in popularity. It was beyond the scope of this thesis to use Hierarchical Bayes and would also have required separate additional software.

In order to use conjoint analysis most effectively it was crucial that a mixture of methods were drawn on at different stages and the findings from these were integrated appropriately (see Figure 4.1 on page 107). This included a systematic literature search, narrative synthesis, cognitive interviews, cross-sectional survey, multinomial logit and latent class analysis. This breadth of methodologies, whilst providing good research training, nevertheless was a challenge. This represented a challenge to the researcher; one of integrating the knowledge gained from multiple methods and attempting to achieve a critical understanding of all of them. However, incorporating traditionally qualitative phenomenon (e.g. GP attitude) into a quantitative framework is one of the benefits of conjoint analysis and so it offered a means of innovating and extending knowledge in the field of determinants of GP consultation.

Equally challenging was the ability to enter into a novel research field (having had no experience of conjoint analysis methods) and attempting to bring the research project into the context of the most recent advances in the field. The ISPOR checklist was helpful in providing a framework for judging the quality of

previous studies (in the systematic review) and developing a conjoint analysis design from literature reviews to developmental qualitative work and understanding their role in the overall project. It is even possible that this study went further than the developers of the ISPOR checklist in using this as a tool for judging and developing studies. However, the ISPOR checklist did not entirely do justice to the more nuanced issues and difficult decisions that were made in the course of the developmental studies. The developmental studies with the Research User Group were crucial at establishing that ACBC (i.e. the most advanced and cutting edge method in the field at the time) was not in fact feasible for this clinical population.

10.4 Implications for future methodological research

Conjoint analysis approaches model hypothetical scenarios as opposed to observed behaviour. This study has gone some way to suggesting that there is scope for investigating whether or not consultation behaviour could be expected to change if service characteristics were to change (e.g. more effective treatments were available). The present study reinforces that stated preference data can be used to estimate parameters for attributes that are not observed in the current health service, or that do not vary in the health service. However, there is great scope for further research to build on the estimates of this study.

A main limitation of conjoint analysis is that it is unclear whether there is a difference between what people say they will do when completing a hypothetical task and what they actually do in the real world. Real world decisions have real consequences. Joint estimation suggests that parameters from the revealed and stated preference data were comparable in one previous study (Mark & Swait

2004), indicating high external validity. However, external validation studies are not common in health research. In the present study the risk of low external validity was minimised by ensuring that during the developmental stages the attributes included in the scenarios were considered as realistic and plausible as possible by the Research User Group. A logical next step to understand external validity better might be to explore the relationship between patients' stated preferences of attributes of the decision to consult the GP for joint pain and their rates of actual GP consultation (revealed preference data). In this way conjoint estimates might be linked directly to healthcare utilisation behaviours (such as consultation) and the external validity of conjoint models assessed, based on large observational epidemiological studies. These studies would need to characterise clinical need and service characteristics. Analyses might be conducted on the combined stated and revealed preference data.

Despite the seeming acceptability of the Partial-Profile Choice-Based Conjoint (74% response) to people with joint pain, this was a highly selected sample and so further work in this clinical population will be required in order to fully test its feasibility/acceptability. Although Partial-Profile Choice-Based Conjoint has limitations and further work is required in the OA population before it can be used more widely, this thesis went some way to testing this method in the area of determinants of GP consultation for joint pain. Partial-Profile Choice-Based Conjoint proved feasible in terms of acceptability to respondents in the form of a postal cross-sectional questionnaire and its ability to quantify the relative importance of service and clinical need factors and the trade-offs between them. It

also ostensibly provides more accurate data on preferences than direct rating methods, or epidemiological and qualitative studies alone.

Future research should continue to look at the potential for using computer-based adaptive designs (as was done in the developmental stages of this study), as increasing numbers of older adults have access to computers and are computer literate. This may open up opportunities for including more complex determinants of consultation without over-burdening respondents. For example, Cunningham et al. (2008), examined preferences for 14 attributes of patient-centred care in hospital service users' from Canadian teaching hospitals using ACBC. These included health information transfer, participation in healthcare decisions and prompt feedback on progress (Cunningham et al. 2008). They concluded that although ACBC methods are more time consuming than traditional conjoint analysis methods, respondents reported ACBC surveys to be more engaging (Cunningham et al. 2010).

The question of whether preferences exist prior to the choice task or whether they change throughout the task as a process of respondents learning and becoming familiar with their own preferences (Cheraghi-Sohi et al. 2007), remains unanswered. Future research could conduct a 'think aloud' interview alongside respondents as they move through the choice tasks from beginning to end. For example, future research could compare the latest software Adaptive Choice-Based Conjoint (ACBC) (designed to adapt and help respondents narrow down their preferences) and a paper based postal Partial-Profile Choice-Based Conjoint survey. This might help to understand how adaptive software (and increased engagement) may help respondents to learn about their preferences

and lead to more precise estimates of preferences. Although the software developers argue that ACBC captures many of the same advantages of Partial-Profile Choice-Based Conjoint, but in a more realistic full-profile context (Orme 2010), this needs further validation.

Future research should validate the claim that by deleting 'irrational' and 'inconsistent' responses there will be greater bias and lower statistical efficiency in conjoint estimates (Lancsar & Louviere 2006). New studies could look at the way in which conjoint analysis can be used in a clinical setting as a means of enabling respondents to become more conscious and aware of their own preferences (as well as the researchers learning about respondents preferences). This would be an interesting application in an area like health psychology where people make choices that are not traditionally considered to be rational. For example, the decision to give up smoking or use drugs (Flach & Diener 2004). Future research could use 'think aloud' techniques alongside the latest adaptive conjoint analysis methods (such as ACBC) to look at the decision-making heuristics used by drug dependents. For example, based on the dual process theories (Stanovich & West 2000) emerges the question of whether alcohol dependents are more likely to use system 1 (intuitive and experiential based decision-making) or system 2 (a systematic process)? Conjoint might assist in identifying the highly individualised determinants of drug use, which might be used to assist in tailoring treatment plans for addiction recovery (GROUP 1998).

A significant limitation of the use of conjoint analysis methods is the need for Sawtooth Software (Inc. Orem, UT) in order to create the partial-profile design. The design and analysis of the design requires both the software and training on

how to use it. The former is expensive and the latter is not easily available in the UK.

10.5 Implications for future clinical practice and research

A prominent clinical implication of this study is that GP consultation for joint pain depends on whether patients believe their GP has a legitimising attitude towards their joint pain. The underlying assumption is that by increasing consultation in general practice this may lead to earlier intervention (including lifestyle advice for weight loss, exercise, physiotherapy and analgesia) thus improving the long-term prognosis for these patients.

Because patients are clearly sensitive towards non-legitimising attitudes and need to feel that their joint pain is taken seriously by their GP, future research would benefit from a deeper understanding of how GPs can be enabled to shift their attitudes towards joint pain. Recent focus groups with GPs suggests that for some GPs a) OA is not a priority, b) that OA treatments are not very efficacious, c) that some perceive that they do not have time to deliver OA care and, d) some are not comfortable with making the diagnosis of OA clinically (without x-rays) (Porcheret: Personal communication). Although it is unclear what proportion of GPs convey a 'normal ageing accept-it' attitude there are clearly some underlying issues which seem to push OA down the priority list for clinicians (at least in a small sample of focus groups). It might be that future clinical research needs to focus on devising messages for GPs to use when discussing/diagnosing OA and provide GP training and education programmes in terms of their role in treating joint pain, or the influence that their attitude may have on a patient's perception towards their joint pain. GP training and education programmes might also target

the way in which other interventions can be used successfully to treat joint pain as well as what is available to patients in terms of chronic pain (OA) management. Future research might draw on some of the work done by May (2004) into how the GPs conceptualisation of chronic illness in the consultation has implications for the delivery of care and how it is experienced by patients (May et al. 2004), the use of empowering explanations and giving explanations to patients that are blame-free (Dowrick 2010).

Equally important is shifting the patients' perception of joint pain as something that is not legitimate and cannot be treated. Although it is perhaps true that the GPs attitude reinforces patient normalising and deprioritising, to some degree shifting this attitude needs to start with the patients themselves. It is often the case that shifting illness perceptions and beliefs is a key aspect of behavioural/psychological interventions which seek to assist patients in managing chronic pain as a long-term condition. For example, a main principle of Acceptance Commitment Therapy (ACT) in chronic pain is that if patients can give up the struggle with pain elimination and stop chasing treatments this will lead them to greater acceptance (McCracken et al. 2005). Future qualitative research should investigate how patients feel about the move to OA being managed as a long-term condition.

In the future studies should include a greater use of conjoint methods to ensure that patients' voice and preferences are being heard (Bridges et al. 2008). This might be achieved by the introduction of the reconfiguration of services and new treatments for osteoarthritis care, using conjoint analysis.

There is potential value for conjoint analysis in exploring the preferences of patients for alternative service model developments within the NHS. Similar to the Caldow et al. (2007) study identified in the systematic review which investigated patient opinion about the provision of nurse-led vs doctor-led primary healthcare in the treatment of minor illness, alternative models of service provision for joint pain specifically should be investigated. For example, in the NHS it is not clear whether patients should consult with GPs for all consultations. There is an increasing shift to other models of service provision such as direct access to non-GPs (e.g. physiotherapists), and practice nurse-led care. Segal et al. (2004) for example, economically evaluated a number of different interventions for the treatment of OA and this included 'primary care: GP, or practise nurse educator plus phone support' (Segal et al. 2004). A logical next step would be to compare joint pain patients' preferences for different OA management provision from a GP, physiotherapist and (nurse-led) telephone support. In addition to healthcare providers, attributes such as length of consultation and treatments (lifestyle/self-management advice) should be included to see at which point a telephone consultation with a healthcare provider becomes acceptable for this type of consultation. The result from this study might extend self-management for OA into the area of tele-health and new information technologies– thus offering a (potentially cost-effective) solution to supporting patients with self-management for OA. For example, online advice services for people with OA is an area that is currently being developed. Choice-based conjoint choice tasks could be used to identify treatment needs and preferences for OA patients online.

In terms of developing new treatments conjoint analysis should be used to

understand in more detail the costs and benefits (trade-offs) of new pharmacological treatments. For example, theories of decision-making as a dual process (Stanovich & West 2000) have been applied to risky decision-making in health (Gibbons et al. 2009) and patient treatment choices (Peters et al. 2007).

According to dual process research patients tend to simplify treatment decisions by avoiding treatments with high risk (Kahneman & Sugden 2005). In an application of ACA in osteoarthrtitis treatments it was certainly evident that older patients were willing to trade-off treatment effectiveness for a lower risk of adverse effects (Fraenkel et al. 2004a). A method such as ACBC (which simplifies decision-making tasks) might be more appropriate for understanding the high risk treatment decisions within osteoarthritis care – if it is possible to develop this method so that it is accessible and user friendly to this clinical population.

Equally useful, conjoint analysis could be used to analyse the factors which influence treatment in terms of the clinical decision-making by health professionals, including the provision of sophisticated imaging like MRI and ultrasound to improve earlier diagnosis of OA (Bijlsma et al. 2011).

The emergence of new treatments (e.g. strontium (Cooper et al. 2011), anti-nerve growth factor (Lane et al. 2010)) and models of care for osteoarthritis, present ongoing opportunities to involve patients and their preferences. This thesis illustrates why conjoint analysis techniques may be useful in this context but also why respondent burden and the steps necessary for rigorous design, analysis and interpretation should not be under-estimated.

References

Adamowicz W, Louviere J, Williams M. 1994. Combining stated and revealed preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management,* 26(3), 271-292.

Alami S, Boutron I, Desjeux D, Hirschhorn M, Meric G, Rannou F, Poiraudeau S. 2011. Patients' and practitioners' views of knee osteoarthritis and its management: a qualitative interview study. *PloS One,* 6(5).

Altman DG. 1991. Practical statistics for medical research. Chapman & Hall. London.

Andersen RM. 1995. Revisiting the behavioral model and access to medical care: does it matter? *Journal of Health and Social Behavior,* 36(3), 1-10.

Arana JE, Leon C, Hanemann MW. 2008. Emotions and decision rules in discrete choice experiments for valuing health care programmes for the elderly. *Journal of Health Economics,* 27(3), 753-769.

Bedson J, Mottram S, Thomas E, Peat G. 2007. Knee pain and osteoarthritis in the general population: what influences patients to consult? *Family Practice,* 24(5), 443-453.

Berry LL, Seiders K, Wilder SS. 2003. Innovations in access to care: a patient-centred approach. *Annals of Internal Medicine,* 139(7), 568-574.

Berry PH, Chapman CR, Covington EC, Dahl J, Katz J, 2001. Pain: Current understanding of assessment, management and treatments. National Pharmaceutical Council and the Joint Commission for the Accreditation of Healthcare Organizations, 29.

Beusterien KM, Dziekan K, Flood E, Harding G, Jordan JC. 2005. Understanding patient preferences for HIV medications using adaptive conjoint analysis: feasibility assessment. *Value in Health,* 8(4), 453-461.

Bijlsma JWJ, Berenbaum F, Lafeber FPJG. 2011. Osteoarthritis: an update with relevance for clinical practice. *The Lancet,* 377(9783), 2115-2126.

Birrell F, Johnell O, Silman A. 1999. Projecting the need for hip replacement over the next three decades: influence of changing demography and threshold for surgery. *Annals of the Rheumatic Diseases,* 58(9), 569-572.

Blagojevic M, Jinks C, Jordan KP. 2008. The influence of consulting primary care on knee pain in older people: a prospective cohort study. *Annals of the Rheumatic Diseases,* 67(12), 1702-1709.

Bowling A. 2005. Quantitative social science: the survey. In: Bowling A, Ebrahim S. Handbook of health research methods: investigation, measurement and analysis. Maidenhead Open University Press, 190-215.

Bradley EH, McGraw SA, Curry L, Buckser A, King KL, Kasl SV, Andersen R. 2002. Expanding the Andersen model: the role of psychosocial factors in long-term care use. *Health Services Research,* 37(5), 1221-1242.

Bradshaw J. 1994. The conceptualization and measurement of need: a social policy perspective. In: Popay J, Williams G. Researching the People's Health. London: Routledge, 45-57.

Brett J, Staniszewska S, Mockford C, Seers K, Herron-Marx S and Bayliss H. 2010. The PIRICOM Study: A systematic review of the conceptualisation, measurement, impact and outcomes of patients and public involvement in health and social care research. London: UKCRC.

Bridges J, Hauber A, Marshall D, Lloyd A, Prosser L, Regier D, Johnson F, Mauskopf J. 2009. A checklist for conjoint analysis applications in health: report of the ISPOR conjoint analysis good research practices taskforce. International Society for Pharmacoeconomics and Outcomes Research. USA.

Bridges JFP, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, Johnson FR, Mauskopf J. 2011. Conjoint analysis applications in health--a checklist: a report of the ISPOR good research practices for conjoint analysis task force. *Value in Health,* 14(4), 403-413.

Bridges JFP, Kinter ET, Kidane L, Heinzen RR, McCormick C. 2008. Things are looking up since we started listening to patients: trends in the application of conjoint analysis in health 19822007. *The Patient: Patient-Centered Outcomes Research,* 1(4), 273-282.

Bridges J, Latalille A, Buttorff C, White S and Niparko J. 2010. A comparison of rating and conjoint analysis methods to measure consumer preferences for hearing aid attributes. The 32nd annual meeting of the society for Medical Decision Making. October 24, 2010.

Broadbent E, Petrie KJ, Main J, Weinman J. 2006. The brief illness perception questionnaire. *Journal of Psychosomatic Research,* 60(6), 631-637.

Caldow J, Bond C, Ryan M, Campbell NC, Miguel FS, Kiger A, Lee A. 2007. Treatment of minor illness in primary care: a national survey of patient satisfaction, attitudes and preferences regarding a wider nursing role. *Health Expectations,* 10(1), 30-45.

Campanelli P. 1997. Testing survey questions: new directions in cognitive interviewing. *Bulletin de Methodologie Sociologique,* 55(1), 5-17.

Campbell SM, Roland MO. 1996. Why do people consult the doctor? *Family Practice,* 13(1), 75-83.

Carroll N, Gagnon J. 1984. Consumer demand for patient-oriented pharmacy services. *American Journal of Public Health,* 74(6), 609-611.

Caruso EM, Rahnev D, Banaji MR. 2009. Using conjoint analysis to detect discrimination: revealing covert preferences from overt choices. *Social Cognition,* 27(1), 128-137.

Chapman CN, Alford JL, Johnson C, Lahav M and Weidemann R. 2009. Comparing results of CBC and ACBC with real product selection. Sawtooth Software Conference. March 23-27.

Chapple A, Campbell S, Rogers A, Roland M. 2002. Users' understanding of medical knowledge in general practice. *Social Science & Medicine,* 54(8), 1215-1224.

Cheraghi-Sohi S, Hole AR, Mead N, McDonald R, Whalley D, Bower P, Roland M. 2008. What patients want from primary care consultations: a discrete choice experiment to identify patients' priorities. *Annals of Family Medicine,* 6(2), 107-115.

Cheraghi-Sohi S, Bower P, Mead N, McDonald R, Whalley D, Roland M. 2007. Making sense of patient priorities: applying discrete choice methods in primary care using 'think aloud' technique. *Family Practice,* 24(3), 276-282.

Coast J, Horrocks S. 2007. Developing attributes and levels for discrete choice experiments using qualitative methods. *Journal of Health Services Research & Policy,* 12(1), 25-30.

Cooper C, Reginster JY, Chapurlat R, Christiansen C, Genant H, Bellamy N, Bensen W, Navarro F, Badurski J, Nasonov E. 2011. Efficacy and safety of oral strontium ranelate for the treatment of knee osteoarthritis: rationale and design of randomised, double-blind, placebo-controlled trial. *Current Medical Research and Opinion,* 28(2), 231-239.

Cronan TA, Shaw WS, Gallagher RA, Weisman M. 1995. Predicting health care use among older osteoarthritis patients in an HMO. *Arthritis & Rheumatism,* 8(2), 66-72.

Cunningham CE, Deal K, Chen Y. 2010. Adaptive Choice-Based Conjoint Analysis: a new patient-centred approach to the assessment of health service preferences. *The Patient: Patient-Centered Outcomes Research,* 3(4), 257-273.

Cunningham CE, Deal K, Rimas H, Campbell H, Russell A, Henderson J, Matheson A, Melnick B. 2008. Using conjoint analysis to model the preferences of different patient segments for attributes of patient-centred care. *The Patient: Patient-Centered Outcomes Research,* 1(4), 317-330.

Cunningham CE, Deal K, Neville A, Rimas H, Lohfeld L. 2006. Modeling the problem-based learning preferences of McMaster University undergraduate medical students using a discrete choice conjoint experiment. *Advances in Health Sciences Education: Theory and Practice,* 11(3), 245-266.

Cunningham CE, Deal K, Rimas H, Buchanan DH, Gold M, Sdao-Jarvie K, Boyle M. 2008. Modeling the information preferences of parents of children with mental health problems: a discrete choice conjoint experiment. *Journal of Abnormal Child Psychology,* 36(7), 1123-1138.

Cunningham CE, Deal K, Rimas H, Chen Y, Buchanan DH, Sdao-Jarvie K. 2009. Providing information to parents of children with mental health problems: a discrete choice conjoint analysis of professional preferences. *Journal of Abnormal Child Psychology,* 37(8), 1089-1102.

Darmon RY, Rouzies D. 1994. Reliability and internal validity of conjoint estimated utility functions under error-free versus error-full conditions. *International Journal of Research in Marketing,* 11(5), 465-476.

de Boer AGEM, Wijker W, de Haes HCJM. 1997. Predictors of health care utilisation in the chronically ill: a review of the literature. *Health Policy,* 42(2), 101-115.

de Vet HCW, Heymans MW, Dunn KM, Pope DP, van der Beek AJ, Macfarlane GJ, Bouter LM, Croft PR. 2002. Episodes of low back pain: a proposal for uniform definitions to be used in research. *Spine,* 27(21), 2409-2416.

Department of Health. 2006. The Musculoskeletal Services Framework. a joint responsibility: doing it differently. Great Britain: Department of Health.

Deyle GD, Henderson NE, Matekel RL, Ryder MG, Garber MB, Allison SC. 2000. Effectiveness of manual physical therapy and exercise in osteoarthritis of the knee. *Annals of Internal Medicine,* 132(3), 173-181.

Dieppe P. 2005. Disease modification in osteoarthritis: Are drugs the answer? *Arthritis & Rheumatism,* 52(7), 1956-1959.

Dieppe P, Basler HD, Chard J, Croft P, Dixon J, Hurley M, Lohmander S, Raspe H. 1999. Knee replacement surgery for osteoarthritis: effectiveness, practice variations, indications and possible determinants of utilisation. *Journal of Rheumatology,* 38(1), 73-83.

Dominick KL, Ahern FM, Gold CH, Heller DA. 2004. Health-related quality of life and health service use among older adults with osteoarthritis. *Arthritis Care & Research,* 51(3), 326-331.

Donovan JL, Blake DR. 2000. Qualitative study of interpretation of reassurance among patients attending rheumatology clinics: “just a touch of arthritis, doctor?”. *British Medical Journal,* 320(7234), 541-544.

Dowrick C. 2010. Medically unexplained symptoms in primary care: how can doctors help, not hinder? *Mental Health in Family Medicine,* 7(4), 191-192.

Dunn KM, Jordan K, Lacey RJ, Shapley M, Jinks C. 2004. Patterns of consent in epidemiologic research: evidence from over 25,000 responders. *American Journal of Epidemiology,* 159(11), 1087-1094.

Dunn KM, Jordan K, Croft PR. 2006. Characterizing the course of low back pain: a latent class analysis. American Journal of Epidemiology,163(8), 754-761.

Edwards P, Roberts I, Clarke M, DiGuiseppi C, Pratap S, Wentz R, Kwan I. 2002. Increasing response rates to postal questionnaires: systematic review. *British Medical Journal,* 324(7347), 1183-1185.

Egnew TR. 2009. Suffering, meaning, and healing: challenges of contemporary medicine. *Annals of Family Medicine,* 7(2), 170-175.

Elrod T, Louviere JJ, Davey KS. 1992. An empirical comparison of ratings-based and choice-based conjoint models. *Journal of Marketing Research,* 29(3), 368-377.

Ethgen O, Kahler KH, Kong SX, Reginster JY, Wolfe F. 2002. The effect of health related quality of life on reported use of health care resources in patients with osteoarthritis and rheumatoid arthritis: a longitudinal analysis. *Journal of Rheumatology,* 29(6), 1147-1155.

Fiebig DG, Haas M, Hossain I, Street DJ, Viney R. 2009. Decisions about pap tests: what influences women and providers? *Social Science & Medicine,* 68(10), 1766-1774.

Flach SD, Diener A. 2004. Eliciting patients' preferences for cigarette and alcohol cessation: an application of conjoint analysis. *Addictive Behaviors,* 29(4), 791-799.

Flynn TN, Louviere JJ, Peters TJ, Coast J. 2007. Best-worst scaling: what it can do for health care research and how to do it. *Journal of Health Economics,* 26(1), 171-189.

Fraenkel L, Bogardus Jr ST, Concato J, Wittink DR. 2004a. Treatment options in knee osteoarthritis: the patient's perspective. *Archives of Internal Medicine,* 164(12), 1299-1304.

Fraenkel L, Wittink DR, Concato J, Fried T. 2004b. Informed choice and the widespread use of antiinflammatory drugs. *Arthritis Care & Research,* 51(2), 210-214.

Fraenkel L, Fried T. 2008. If you want patients with knee osteoarthritis (OA) to exercise: tell them about NSAIDS. *The Patient,* 1(1), 21-26.

Gerard K, Salisbury C, Street D, Pope C, Baxter H. 2008. Is fast access to general practice all that should matter? A discrete choice experiment of patients' preferences. *Journal of Health Services Research & Policy,* 13, 3-10.

Gibbons FX, Houlihan AE, Gerrard M. 2009. Reason and reaction: the utility of a dual-focus, dual-processing perspective on promotion and prevention of adolescent health risk behaviour. *British Journal of Health Psychology,* 14(2), 231-248.

Gignac MAM, Davis AM, Hawker G, Wright JG, Mahomed N, Fortin PR, Badley EM. 2006. “What do you expect? You're just getting older”: a comparison of perceived osteoarthritis-related and aging-related health experiences in middle-and older-age adults. *Arthritis Care & Research,* 55(6), 905-912.

Gilbride TJ, Allenby GM. 2004. A choice model with conjunctive, disjunctive, and compensatory screening rules. *Marketing Science,* 23(3) 391-406.

Gluud LL. 2006. Bias in clinical intervention research. *American Journal of Epidemiology,* 163(6), 493-501.

Graf MA, Tanner DD, Swinyard WR. 1993. Optimising the delivery of patient and physician satisfaction: a conjoint analysis approach. *Healthcare Management Review,* 18(4), 34-43.

Green LA, Fryer GE, Yawn BP, Lanier D, Dovey SM. 2001. The ecology of medical care revisited. *New England Journal of Medicine,* 344(26), 2021-2025.

Green PE, Rao VR. 1971. Conjoint measurement for quantifying judgmental data. *Journal of Marketing Research,* 8, 355-363.

Grol R, De Maeseneer J, Whitfield M, Mokkink H. 1990. Disease-centred versus patient-centred attitudes: comparison of general practitioners in Belgium, Britain and The Netherlands. *Family practice,* 7(2), 100-103.

Group PMR. 1998. Matching patients with alcohol disorders to treatments: clinical implications from project MATCH. *Journal of Mental Health,* 7(6), 589-602.

Haas M. 2005. The impact of non-health attributes of care on patients' choice of GP. *Australian Journal of Primary Health,* 11(1), 40-46.

Hagen KB, Bjørndal A, Uhlig T, Kvien TK, 2000. A population study of factors associated with general practitioner consultation for non-inflammatory musculoskeletal pain. *Annals of the Rheumatic Diseases,* 59(10), 788-793.

Haider W, Ewing GO. 1990. A model of tourist choices of hypothetical Caribbean destinations. *Leisure Sciences,* 12(1), 33-47.

Hale E, Treharne G, Kitas G. 2007. The common-sense model of self-regulation of health and illness: how can we use it to understand and respond to our patients' needs? *Journal of Rheumatology,* 46(6), 904-906.

Hartvigsen J, Christensen K. 2007. Active lifestyle protects against incident low back pain in seniors: a population-based 2-year prospective study of 1387 Danish twins aged 70–100 years. *Spine,* 32(1), 76.

Harwood RH, Rogers A, Dickinson E, Ebrahim S. 1994. Measuring handicap: the London Handicap Scale, a new outcome measure for chronic disease. *Quality in Health Care,* 3(1), 11-16.

Hawker G, French M, Elkayam J, Davis A. 2010. Predictability of intermittent hip/knee OA pain. *Osteoarthritis and Cartilage,* 18, 157.

Hawker G, Stewart L, French M, Cibere J, Jordan J, March L, Suarez-Almazor M, Gooberman-Hill R. 2008. Understanding the pain experience in hip and knee osteoarthritis–an OARSI/OMERACT initiative. *Osteoarthritis and Cartilage,* 16(4), 415-422.

Hensher DA, Rose JM, Greene WH. 2005. Applied choice analysis: a primer. Cambridge University Press.

Higgins JPT, Green S, Collaboration C. 2008. Cochrane handbook for systematic reviews of interventions. John Wiley and Sons Ltd. London.

Hill S, Dziedzic K, Thomas E, Baker SR, Croft P. 2007. The illness perceptions associated with health and behavioural outcomes in people with musculoskeletal hand problems: findings from the North Staffordshire Osteoarthritis Project (NorStOP). *Journal of Rheumatology,* 46(6), 944-951.

Hjelmgren J, Anell A. 2007. Population preferences and choice of primary care models: a discrete choice experiment in Sweden. *Health Policy,* 83(2-3), 314-322.

Hopman-Rock M, de Bock, G.H, Bijlsma JW, Springer MP, Hofman A, Kraaimaat FW. 1997. The pattern of health care utilisation of elderly people with arthritic pain in the hip or knee. *The International Society for Quality in Health Care,* 9(2), 129-137.

Huber J. 2005. Conjoint analysis: how we got here and where we are (an update). *Sawtooth Software Research Paper Series, Duke University, Durham, North Carolina.*

Huber, J, Zwerina K. 1996. The importance of utility balance in efficient choice designs. J*ournal of Marketing Research*, 33 (3), 307–317.

Huber M, Knottnerus JA, Green L, Horst H, Jadad AR, Kromhout D, Leonard B, Lorig K, Loureiro MI, Meer JWM. 2011. How should we define health? *British Medical Journal,* 343.

Hundley V, Ryan M. 2004. Are women's expectations and preferences for intrapartum care affected by the model of care on offer? *BJOG: An International Journal of Obstetrics and Gynaecology,* 111(6), 550-560.

International Association for the Study of Pain, 2010-last update, The declaration of Montreal.Available:http://www.iasppain.org/Content/NavigationMenu/Advocacy/DeclarationofMontr233al/default.htm [22nd January, 2012].

Ives J, Damery S, Redwod S. 2012. PPI, paradoxes and Plato: who's sailing the ship? Online advance: doi:10.1136/medethics-2011-100150. *Journal* of Medical Ethics.

Jensen MP, Keefe FJ, Lefebvre JC, Romano JM, Turner JA. 2003. One-and two-item measures of pain beliefs and coping strategies. *Pain,* 104(3), 453-469.

Jinks C, Jordan K, Ong BN, Croft P. 2004. A brief screening tool for knee pain in primary care (KNEST). 2. Results from a survey in the general population aged 50 and over. *Journal of Rheumatology,* 43(1), 55-61.

Jinks C, Ong BN, Richardson J. 2007. A mixed methods study to investigate needs assessment for knee pain and disability: population and individual perspectives. *BMC Musculoskeletal Disorders,* 8(59).

Johnson FR. 2008. Why not ask?: Measuring patient preferences for healthcare decision making. *The Patient: Patient-Centered Outcomes Research,* 1(4), 245-248.

Johnson RM and Orme BK. 2007. A new approach to Adaptive CBC. Sawtooth Software Conference Proceedings, Sequim, WA.

Johnson RM and Orme BK. 1996. How many questions should you ask in choice-based conjoint studies. Conference Proceedings of the ART Forum.

Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. 2010. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. *BMC Musculoskeletal Disorders,* 11(144).

Jordan K, Clarke AM, Symmons DPM, Fleming D, Porcheret M, Kadam UT, Croft P. 2007. Measuring disease prevalence: a comparison of musculoskeletal disease using four

general practice consultation databases. *The British Journal of General Practice: Royal College of General Practitioners,* 57(534), 7-14.

Jordan K, Jinks C, Croft P. 2006. A prospective study of the consulting behaviour of older people with knee pain. *British Journal of General Practice,* 56(525), 269-276.

Kadam UT, Croft PR. 2007. Clinical comorbidity in osteoarthritis: associations with physical function in older patients in family practice. *Journal of Rheumatology,* 34(9), 1899-1904.

Kahneman D, Sugden R. 2005. Experienced utility as a standard of policy evaluation. *Environmental and Resource Economics,* 32(1), 161-181.

Kee CC. 1998. Living with osteoarthritis: Insiders' views. *Applied Nursing Research,* 11(1), 19-26.

Kessels R, Goos P, Vandebroek M. 2006. A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research,* 43(8), 409-419.

Kievit W, van Hulst L, van Riel P, Fraenkel L. 2010. Factors that influence rheumatologists' decision to escalate care in RA: results from a choice based conjoint analysis. *Arthritis Care and Research,* 62(6), 842-847.

Kim MY, Park JK, Koh SB, Kim CB. 2010. Factors influencing utilisation of medical care among osteoarthritis patients in Korea: using 2005 Korean National Health and Nutrition Survey Data. *Journal of Preventive Medicine and Public Health,* 43(6), 513-522.

Kish L. 1965. Survey sampling. John Wiley and Sons Ltd. New York.

Lancsar E, Louviere J. 2006. Deleting 'irrational' responses from discrete choice experiments: a case of investigating or imposing preferences? *Journal of Health Economics,* 15(8), 797-811.

Lancsar E, Louviere J. 2008. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. *PharmacoEconomics,* 26(8), 661-677.

Lane NE, Schnitzer TJ, Birbara CA, Mokhtarani M, Shelton DL, Smith MD, Brown MT. 2010. Tanezumab for the treatment of pain from osteoarthritis of the knee. *New England Journal of Medicine,* 363(16), 1521-1531.

Li LC, Sayre EC, Kopec JA, Esdaile JM, Bar S, Cibere J. 2011. Quality of nonpharmacological care in the community for people with knee and hip osteoarthritis. *Journal of Rheumatology,* 38(10), 2230-2237.

Loeser JD. 2000. Pain and suffering. *Clinical Journal of Pain,* 16(2), 2-6.

Longo MF, Cohen DR, Hood K, Edwards A, Robling M, Elwyn G, Russell IT. 2006. Involving patients in primary care consultations: assessing preferences using discrete choice experiments. *The British Journal of General Practice,* 56(522), 35-42.

Lorig KR, Sobel DS, Stewart AL, Brown Jr BW, Bandura A, Ritter P, Gonzalez VM, Laurent DD, Holman HR. 1999. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial. *Medical Care,* 37(1), 5.

Louviere J, Flynn T, Carson R. 2010. Discrete choice experiments are not conjoint analysis. *Journal of Choice Modelling,* 3(3), 57-72.

Louviere JJ. 1988. Analyzing decision making: metric conjoint analysis. Sage University paper series on Quantitative applications in the social sciences. California.

Louviere JJ, Hensher DA, Swait JD. 2000. Stated choice methods: analysis and applications. Cambridge University Press.

Louviere JJ, Lancsar E. 2009. Choice experiments in health: the good, the bad, the ugly and toward a brighter future. *Journal of Health Economics, Policy and Law,* 4, 527-546.

Luce RD, Tukey JW. 1964. Simultaneous conjoint measurement: a new type of fundamental measurement. *Journal of Mathematical Psychology,* 1(1), 1-27.

Mallen CD, Peat G. 2009. Discussing prognosis with older people with musculoskeletal pain: a cross-sectional study in general practice. *BMC Family Practice,* 10, 50-50.

Mann C, Gooberman-Hill R. 2011. Healthcare provision for osteoarthritis: concordance between what patients would like and what health professionals think they should have. *Arthritis Care & Research,* 63(7), 963-972.

Mant D. 1998. R&D in primary care: an NHS priority. *British Journal of General Practice,* 48(426), 871.

Mariampolski H. 2001. Qualitative market research: a comprehensive guide. Sage Publications, Inc. London.

Mark TL, Swait J. 2004. Using stated preference and revealed preference modeling to evaluate prescribing decisions. *Journal of Health Economics,* 13(6), 563-573.

Markham FW, Diamond JJ, Hermansen CL. 1999. The use of conjoint analysis to study patient satisfaction. *Evaluation & the Health Professions,* 22(3), 371-378.

Marshall D, Bridges JFP, Hauber B, Cameron R, Donnalley L, Fyie K, Johnson FR. 2010. Conjoint analysis applications in health how are studies being designed and reported?: an update on current practice in the published literature between 2005 and 2008. *The Patient: Patient-Centered Outcomes Research,* 3(4), 249-256.

Martin C. 2005. Epidemiological study designs for health care research and evaluation. In: Bowling A, Ebrahim S. Handbook of health research methods: investigation, measurement and analysis. Maidenhead Open University Press, 98-164.

Mathers CD, Stein C, Fat DM, Rao C, Inoue M, Tomijima N, Bernard C, Lopez AD, Murray CJL. 2002. Global burden of disease 2000: version 2, methods and results. *Geneva: WHO.*

May C, Allison G, Chapple A, Chew-Graham C, Dixon C, Gask L, Graham R, Rogers A, Roland M. 2004. Framing the doctor-patient relationship in chronic illness: a comparative study of general practitioners' accounts. *Sociology of Health & Illness,* 26(2), 135-158.

McClain JO, Rao VR. 1974. Trade-offs and conflicts in evaluation of health system alternatives: methodology for analysis. *Health Services Research,* 9(1), 35-52.

McCormick A, Fleming D, Charlton J. 1995. Morbidity statistics from general practice fourth national study 1991-1992. Office of Population Censuses and Surveys. London.

McCracken LM, Eccleston C, Bell L. 2005. Clinical assessment of behavioral coping responses: preliminary results from a brief inventory. *European Journal of Pain,* 9(1), 69-78.

McDonald DD, Shea M, Fedo J, Rose L, Bacon K, Noble K, Stewart J. 2008. Older adult pain communication and the Brief Pain Inventory Short Form. *Pain Management Nursing: Official Journal of the American Society of Pain Management Nurses,* 9(4), 154-159.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. Frontiers in Econometrics. New York Academic Press, 105-142.

Mead N, Bower P. 2000. Patient-centredness: a conceptual framework and review of the empirical literature. *Social Science & Medicine,* 51(7), 1087-1110.

Miller GA. 1956. The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review,* 63(2), 81-97.

Mitchell HL, Carr AJ, Scott DL. 2006. The management of knee pain in primary care: factors associated with consulting the GP and referrals to secondary care. *Journal of Rheumatology,* 45(6), 771-776.

Moher D, Liberati A, Tetzlaff J, Altman DG. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine,* 6(7), 1371.

Moore WL, Gray-Lee J, Louviere JJ. 1998. A cross-validity comparison of conjoint analysis and choice models at different levels of aggregation. *Marketing Letters,* 9(2), 195-207.

Morgan A, Shackley P, Pickin M, Brazier J. 2000. Quantifying patient preferences for out-of-hours primary care. *Journal of Health Services Research & Policy,* 5(4), 214-218.

Murphy LB, Helmick CG, Schwartz TA, Renner JB, Tudor G, Koch GG, Dragomir AD, Kalsbeek WD, Luta G, Jordan JM. 2010. One in four people may develop symptomatic hip osteoarthritis in his or her lifetime. *Osteoarthritis and Cartilage,* 18(11), 1372-1379.

National collaborating centre for chronic conditions. 2008. Osteoarthritis: national clinical guideline for care and management in adults. London: Royal College of Physicians.

Nickerson CAE, McClelland GH, Petersen DM. 1991. Measuring contraceptive values: an alternative approach. *Journal of Behavioral Medicine,* 14(3), 241-266.

Office of National Statistics. 1996. Office of National Statistics; 1996 based national population projections. London: The Stationery Office.

Ong B, Jordan K, Richardson J, Croft P. 1999. Experiencing limiting long-standing illness (short report). *Health & Social Care in the Community,* 7(1), 61–68.

Orme BK, Alpert MI and Christensen E. 1997. Assessing the validity of conjoint analysis–continued. Sawtooth Software Conference Proceedings. pp 209-225.

Orme BK. 2006. Getting started with conjoint analysis: strategies for product design and pricing research. Research Publishers, LLC. Madison.

Orme BK. 2010. SSI Web v7.0 Software for web interviewing and conjoint analysis. Sawtooth Software.

Parker BR, Srinivasan V. 1976. A consumer preference approach to the planning of rural primary health-care facilities. *Operations Research,* 24(5), 991-1025.

Patterson M and Chrzan K. 2004. Partial profile discrete choice: what's the optimal number of attributes? Sawtooth Software Conference Proceedings. October 6-8 pp173–185.

Payne JW, Bettman JR, Schkade DA. 1999. Measuring constructed preferences: towards a building code. *Journal of Risk and Uncertainty,* 19(1), 243-270.

Peacock S, Apicella C, Andrews L, Tucker K, Bankier A, Daly M, Hopper J. 2006. A discrete choice experiment of preferences for genetic counselling among Jewish women seeking cancer genetics services. *British Journal of Cancer,* 95(10), 1448-1453.

Pearmain D, Swanson J, Kroes E, Bradley M. 1991. Stated preference techniques: a guide to practice. Steer Davies & Gleave Ltd. London.

Peat G, Thomas E. 2009. When knee pain becomes severe: a nested case-control analysis in community-dwelling older adults. *Journal of Pain,* 10(8), 798-808.

Peat G, McCarney R, Croft P. 2001. Knee pain and osteoarthritis in older adults: a review of community burden and current use of primary health care. *Annals of the Rheumatic Diseases,* 60(2), 91-97.

Perruccio AV, Power JD, Badley EM. 2006. Revisiting arthritis prevalence projections: it's more than just the aging of the population. *Journal of Rheumatology,* 33(9), 1856-1862.

Peters E, Hess TM, Västfjäll D, Auman C. 2007. Adult age differences in dual information processes: Implications for the role of affective and deliberative processes in older adults' decision making. *Perspectives on Psychological Science,* 2(1), 1-23.

Phillips KA, Johnson FR, Maddala T. 2002a. Measuring what people value: a comparison of "attitude" and "preference" surveys. *Health Services Research,* 37(6), 1659-1679.

Phillips KA, Maddala T, Johnson FR. 2002b. Measuring preferences for health care interventions using conjoint analysis: an application to HIV testing. *Health Services Research,* 37(6), 1681-1705.

Phillips KA, Morrison KR, Andersen R, Aday LA. 1998. Understanding the context of healthcare utilisation: assessing environmental and provider-related variables in the behavioral model of utilisation. *Health Services Research,* 33(3 Pt 1), 571-596.

Pieterse AH, Stiggelbout AM, Marijnen CAM. 2010. Methodologic evaluation of adaptive conjoint analysis to assess patient preferences: an application in oncology. *Health Expectations,* 13(4), 392-405.

Porcheret M, Jordan K, Jinks C. 2007. Primary care treatment of knee pain—a survey in older adults. *Journal of Rheumatology,* 46(11), 1694 -1700.

Rao J, Callahan L, Helmick C. 1997. Characteristics of persons with self-reported arthritis and other rheumatic conditions who do not see a doctor. *Journal of Rheumatology,* 24(1), 169-173.

Ratcliffe J, Buxton M, McGarry T, Sheldon R, Chancellor J. 2004. Patients' preferences for characteristics associated with treatments for osteoarthritis. *Journal of Rheumatology,* 43(3), 337-345.

Reardon G, Pathak DS. 1990. Segmenting the antihistamine market: an investigation of consumer preferences. *Journal of Health Care Marketing,* 10(3), 23-33.

Roddy E, Zhang W, Doherty M, Arden NK, Barlow J, Birrell F, Carr A, Chakravarty K, Dickson J, Hay E. 2005. Evidence-based recommendations for the role of exercise in the management of osteoarthritis of the hip or knee—the MOVE consensus. *Journal of Rheumatology,* 44(1), 67-73.

Rosemann T, Joos S, Szecsenyi J, Laux G, Wensing M. 2007. Health service utilisation patterns of primary care patients with osteoarthritis. *BMC Health Services Research,* 7 (169).

Rosemann T, Wensing M, Joest K, Backenstrass M, Mahler C, Szecsenyi J. 2006. Problems and needs for improving primary care of osteoarthritis patients: the views of patients, general practitioners and practice nurses. *BMC Musculoskeletal Disorders,* 7 (48).

Rosenthal DA, Chan F, Livneh H. 2006. Rehabilitation students' attitudes toward persons with disabilities in high and low-stakes social contexts: A conjoint analysis. *Disability and Rehabilitation,* 28(24), 1517-1527.

Rosko MD, Walker LR, McKenna W, DeVita M. 1983. Measuring consumer preferences for ambulatory medical care arrangements. *Journal of Medical Systems,* 7(6), 545-554.

Rubin G, Bate A, George A, Shackley P, Hall N. 2006. Preferences for access to the GP: a discrete choice experiment. *The British Journal of General Practice: Journal of the Royal College of General Practitioners,* 56(531), 743-748.

Ryan M. 1999. Using conjoint analysis to take account of patient preferences and go beyond health outcomes: an application to in vitro fertilisation. *Social Science & Medicine,* 48(4), 535-546.

Ryan M, McIntosh E, Shackley P. 1998. Using conjoint analysis to elicit the views of health service users: an application to the patient health card. *Health Expectations,* 1(2), 117-129.

Ryan M, Watson V, Amaya-Amaya M. 2003. Methodological issues in the monetary valuation of benefits in healthcare. *Expert Review of Pharmacoeconomics and Outcomes Research,* 3(6), 717-727.

Ryan M, Farrar S. 2000. Using conjoint analysis to elicit preferences for health care. *British Medical Journal, (Clinical research ed.),* 320(7248), 1530-1533.

Ryan M, Scott DA, Reeves C, Bate A, van Teijlingen , E.R, Russell EM, Napper M, Robb CM. 2001. Eliciting public preferences for healthcare: a systematic review of techniques. *Health Technology Assessment,* 5(5), 1-186.

Ryan M, Watson V, Entwistle V. 2009. Rationalising the 'irrational': a think aloud study of discrete choice experiment responses. *Journal of Health Economics,* 18(3), 321-336.

Ryan M, Gerard K. 2003. Using discrete choice experiments to value health care programmes: current practice and future research reflections. *Applied Health Economics and Health Policy,* 2(1), 55-64.

Ryan M, Major K, Skåtun D. 2005. Using discrete choice experiments to go beyond clinical outcomes when evaluating clinical practice. *Journal of Evaluation in Clinical Practice,* 11(4), 328-338.

San Miguel F. 2000. Testing the assumptions of completeness, stability and rationality of preferences in health economics using discrete choice experiments. University of Aberdeen.

Sanders C, Donovan J, Dieppe P. 2002. The significance and consequences of having painful and disabled joints in older age: co-existing accounts of normal and disrupted biographies. *Sociology of Health & Illness,* 24(2), 227-253.

Sanders C, Donovan J, Dieppe P. 2004. Unmet need for joint replacement: a qualitative investigation of barriers to treatment among individuals with severe pain and disability of the hip and knee. *Journal of Rheumatology,* 43(3), 353-357.

Scott A, Vick S. 1999. Patients, doctors and contracts: an application of principal-agent theory to the doctor-patient relationship. *Scottish Journal of Political Economy,* 46(2), 111-134.

Scott A, Watson MS, Ross S. 2003. Eliciting preferences of the community for out-of-hours care provided by general practitioners: a stated preference discrete choice experiment. *Social Science & Medicine,* 56(4), 803-814.

Segal L, Day S, Chapman A, Osborne R. 2004. Can we reduce disease burden from osteoarthritis? *The Medical Journal of Australia,* 180(5), 11-17.

Singh J, Cuttler L, Shin M, Silvers J, Neuhauser D. 1998. Medical decision-making and the patient: understanding preference patterns for growth hormone therapy using conjoint analysis. *Medical Care,* 36(8), 31-45.

Stanovich KE, West RF. 2000. Individual differences in reasoning: implications for the rationality debate? *Behavioral and Brain Sciences,* 23(5), 645-665.

Steel N, Bachmann M, Maisey S, Shekelle P, Breeze E, Marmot M, Melzer D. 2008. Self-reported receipt of care consistent with 32 quality indicators: national population survey of adults aged 50 or more in England. *British Medical Journal,* 337 (7667) 441-445.

Stevens A, Gabbay J. 1991. Needs assessment. *Health trends,* 23(1), 20-23.

Stiggelbout AM, de Vogel-Voogt E. 2008. Health state utilities: a framework for studying the gap between the imagined and the real. *Value in Health,* 11(1), 76-87.

Stiggelbout AM, Vogel-Voogt E, Noordijk EM, Vliet Vlieland TP. 2008. Individual quality of life: adaptive conjoint analysis as an alternative for direct weighting? *Quality of life Research,* 17(4), 641-649.

Stoller EP, Forster LE, Portugal S. 1993. Self-care responses to symptoms by older people: A health diary study of illness behavior. *Medical Care,* 31(1), 24-42.

Strain LA. 1991. Use of health services in later life: the influence of health beliefs. *Journal of Gerontology,* 46(3), 143-150.

Tallon D, Chard J, Dieppe P. 2000. Exploring the priorities of patients with osteoarthritis of the knee. *Arthritis Care & Research,* 13(5), 312-319.

Terwee CB, Bot SDM, De Boer MR, van der Windt DAWM, Knol DL, Dekker J, Bouter LM, De Vet HCW. 2007. Quality criteria were proposed for measurement properties of health status questionnaires. *Journal of Clinical Epidemiology,* 60(1), 34-42.

Thomas, R. 1999-last update, Income Question bank topic commentary on income [Homepage of (The question bank is an ESRC funded internet social resource held within the department of sociology, university of Surrey)], [Online]. Available: http//qb.soc.surrey.ac.uk/topics/income/thomas_wealth.htm.

Thomas E, Wilkie R, Peat G, Hill S, Dziedzic K, Croft P. 2004. The North Staffordshire Osteoarthritis Project--NorStOP: prospective, 3-year study of the epidemiology and management of clinical osteoarthritis in a general population of older adults. *BMC Musculoskeletal Disorders,* 5, 2-2.

Thorstensson CA, Gooberman-Hill R, Adamson J, Williams S, Dieppe P. 2009. Help-seeking behaviour among people living with chronic hip or knee pain in the community. *BMC Musculoskeletal Disorders,* 7(10), 153-163.

Tourangeau R, Yan T. 2007. Sensitive questions in surveys. *Psychological Bulletin,* 133(5), 859– 883.

Turk DC, Dworkin RH, Burke LB, Gershon R, Rothman M, Scott J, Allen RR, Atkinson JH, Chandler J, Cleeland C, Cowan P, Dimitrova R, Dionne R, Farrar JT, Haythornthwaite JA, Hertz S, Jadad AR, Jensen MP, Kellstein D, Kerns RD, Manning DC, Martin S, Max MB, McDermott MP, McGrath P, Moulin DE, Nurmikko T, Quessy S, Raja S, Rappaport BA, Rauschkolb C, Robinson JP, Royal MA, Simon L, Stauffer JW, Stucki G, Tollett J, von Stein T, Wallace MS, Wernicke J, White RE, Williams AC, Witter J, Wyrwich KW. 2006. Developing patient-reported outcome measures for pain clinical trials: IMMPACT recommendations. *Journal of Pain,* 125(3), 208-215.

Turner D, Tarrant C, Windridge K, Bryan S, Boulton M, Freeman G, Baker R. 2007. Do patients value continuity of care in general practice? An investigation using stated preference discrete choice experiments. *Journal of Health Services Research & Policy,* 12(3), 132-137.

van Til J, Stiggelbout AM, Ijzerman MJ. 2009. The effect of information on preferences stated in a choice-based conjoint analysis. *Patient Education & Counseling,* 74(2), 264-271.

Vick S, Scott A. 1998. Agency in health care. Examining patients' preferences for attributes of the doctor-patient relationship. *Journal of Health Economics,* 17(5), 587-605.

Victor CR, Ross F, Axford J. 2004. Capturing lay perspectives in a randomized control trial of a health promotion intervention for people with osteoarthritis of the knee. *Journal of Evaluation in Clinical Practice,* 10(1), 63-70.

Viney R, Lancsar E, Louviere J. 2002. Discrete choice experiments to measure consumer preferences for health and healthcare. *Expert Review of Pharmacoeconomics and Outcomes Research,* 2(4), 319-326.

Wagner EH, Davis C, Schaefer J, Von Korff M, Austin B, 1999. A survey of leading chronic disease management programs: are they consistent with the literature? *Managed Care Quarterly,* 7(3), 56-66.

Willis GB. 1999-last update, Cognitive interviewing: a “how to” guide [Homepage of Research Triangle Institute], [Online]. Available: fog.its.uiowa.edu/~c07b209/interview.pdf [25/06, 2009].

Wittink DR and Bergestuen T. 2001. Forecasting with conjoint analysis. In: Armstrong JS, ed. Principles of forecasting: a handbook for researchers and practitioners. Kluwer Academic Publishers. Norwell.

Wong DW, Chan F, Da Silva Cardoso E, Lam CS, Miller SM. 2004. Rehabilitation counseling students' attitudes toward people with disabilities in three social contexts: a conjoint analysis. *Rehabilitation counselling bulletin,*47(4), 194-204.

World Health Organization. 2003. The burden of musculoskeletal conditions at the start of the new millennium. 919. WHO technical report series.

Yates JF, Patalano AL. 1999. Decision making and aging. In: Park DC, Morrell RW. Processing of Medical Information in Aging Patients. Mahwah, Lawrence Erlbaum Associates, 31–54.

Zhang W, Doherty M, Peat G, Bierma-Zeinstra M, Arden N, Bresnihan B, Herrero-Beaumont G, Kirschner S, Leeb B, Lohmander L. 2010a. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. *Annals of the Rheumatic Diseases,* 69(3), 483-489.

Zhang W, Moskowitz R, Nuki G, Abramson S, Altman R, Arden N, Bierma-Zeinstra S, Brandt K, Croft P, Doherty M. 2007. OARSI recommendations for the management of hip and knee osteoarthritis, part I: critical appraisal of existing treatment guidelines and systematic review of current research evidence. *Osteoarthritis and Cartilage,* 15(9), 981-1000.

Zhang W, Nuki G, Moskowitz R, Abramson S, Altman R, Arden N, Bierma-Zeinstra S, Brandt K, Croft P, Doherty M. 2010b. OARSI recommendations for the management of hip and knee osteoarthritis: Part III: changes in evidence following systematic cumulative update of research published through January 2009. *Osteoarthritis and Cartilage,* 18(4), 476-499.

Zigmond AS, Snaith R. 1983. The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica,* 67(6), 361-370.

Appendices

Appendix 1a: Oral presentation abstracts

Coxon D, Frisher M, Jinks C, Jordan K, Peat G. Deciding to consult the general practitioner for joint pain: a choice-based conjoint analysis study.

North American Primary Care Research Group. New Orleans, United States of America. December 2012

Context: Symptomatic osteoarthritis (OA) is a major cause of disability. However, even among people with persistent, severe pain, many often do not consult their general practitioner (GP). Previous studies have focussed on patient characteristics and clinical need as determinants of consultation. Yet, given reports of patients' negative experiences and low expectations of healthcare services, their perceptions of what primary care has to offer will also be important.

Objective: To investigate the relative importance of service-related and clinical need attributes in the decision to consult a GP for symptomatic OA.

Design: Partial-profile choice-based conjoint analysis study, cross-sectional survey, single postal self-complete questionnaire.

Setting: General population.

Participants: Adults aged 50 years and over with hip, knee, or hand pain identified from an existing population cohort study.

Instruments: Questionnaire containing 10 choice tasks, each presenting two scenarios based on a combination of three out of six selected attributes (pain characteristics, pain disruption to everyday life, other current health problems, assessment/investigations available, treatment options available, and GP attitude).

Results: 863 (74%) people responded (55% female) (mean age 70 years, range: 58-93). The most important determinant of the patient's decision to consult the GP for joint pain

was the extent to which pain disrupted everyday life (1.10 logits). GP attitude (0.86) was perceived to be more important than the available treatments (0.45) or assessment/investigations (0.48). The decision to consult the GP for joint pain was less influenced by other health problems (0.46) and episodes of more severe, unpredictable pain (0.16).

Conclusions: Service-related factors are as influential as the most important need-related determinants of consultation. Believing the GP would regard joint pain as 'part of the normal ageing process that one just has to accept' is a strong disincentive to seeking help, potentially outweighing other aspects of quality of care.

Coxon D, Frisher M, Jinks C, Jordan K, Peat G. 2012. Deciding to consult the general practitioner for symptomatic osteoarthritis: a choice-based conjoint analysis study, Rheumatology (vol. 51. P. 37) Glasgow, Scotland. 1– 3 May 2012

BACKGROUND: Symptomatic osteoarthritis (OA) is a major cause of disability. However, even among people with persistent, severe pain, many often do not consult their general practitioner (GP) about it over several years. Previous studies of the determinants of consultation for OA have focussed almost exclusively on patient characteristics and clinical need. Yet, given reports of patients' negative experiences and low expectations of healthcare services, their perceptions of what primary care has to offer will also be important. This study investigated the relative importance of service-related factors and clinical need factors in the decision to consult a GP for symptomatic OA.

METHODS: The design was a partial-profile choice-based conjoint analysis study based on a cross-sectional survey using a single postal self-complete questionnaire. Adults aged 50 years and over with hip, knee, or hand pain were identified from an existing population cohort study. Eligible potential participants were sent a postal self-complete questionnaire containing 10 choice tasks, each presenting two scenarios based on a combination of three out of six selected attributes (pain characteristics, pain disruption to everyday life, other current health problems, assessment/investigations available, treatment options available, and GP attitude). Multinomial logit regression (main effects) was used to estimate the relative importance patients placed on each of the attributes. Results represent the

aggregate difference between the maximum and minimum utilities, expressed as logits, for each attribute level. Latent class analysis was used to identify subgroups of respondents with similar preferences.

RESULTS: 863 (74%) people responded (55% female) (mean age 70 years, range: 58-93). The most important determinant of the patient's decision to consult the GP for joint pain was the extent to which pain disrupted everyday life (1.10 logits). GP attitude (0.86) was perceived to be more important than the available treatments (0.45) or assessment/investigations (0.48). The decision to consult the GP for joint pain was less influenced by other health problems (0.46) and episodes of more severe, unpredictable pain (0.16). Subgroups identified by latent class appeared to be based more on differences in the strength of preferences overall than to differences in the relative importance of attributes per se.

CONCLUSIONS: Service-related factors are as influential as the most important need-related determinant of consultation (i.e. pain disruption). Believing the GP would regard joint pain as 'part of the normal ageing process that one just has to accept' is a strong disincentive to seeking help, potentially outweighing other aspects of quality of care. Partial-profile conjoint methods are acceptable to respondents, well-completed, and can address attributes that are less accessible in traditional epidemiological designs.

Coxon D, Frisher M, Jinks C, Jordan K, Peat G. The decision to consult the general practitioner for joint pain: a choice-based conjoint analysis study. Society for Academic Primary Care. Kendal, England. 24– 25 November 2011

SUMMARY: Symptomatic osteoarthritis (OA) is a leading cause of years lived with disability. However, even among people with persistent, severe pain, many appear not to consult their general practitioner (GP) about it for years at a time. Previous studies of the determinants of consultation for OA have focussed almost exclusively on patient characteristics and clinical need. Yet, given reports of patients' negative experiences and low expectations of healthcare services, their perceptions of what primary care has to offer may also be important. This conjoint study investigated the relative importance of selected service-related factors alongside clinical need.

METHODS: The design was a partial-profile choice-based conjoint analysis study based on a cross-sectional survey using a single postal self-complete questionnaire. Adults aged 50 years and over with hip, knee, or hand pain were identified from an existing population cohort study. Eligible potential participants were sent a postal self-complete questionnaire containing 10 choice tasks, each presenting two scenarios based on a combination of three out of six selected attributes (pain characteristics, pain disruption to everyday life, other current health problems, assessment/investigations available, treatment options available, and GP attitude). Multinomial logit regression (main effects) was used to estimate the relative importance patients placed on each of the attributes. Results were

summarised as the aggregate difference between the maximum and minimum utilities, expressed as logits, for levels within each attribute.

FINDINGS: Analyses were based on 863 respondents (55% female, mean age 70 years (range: 58-93) providing complete data to all choice tasks (74% response). The most important determinant of the patient's decision to consult the GP for joint pain was the extent to which pain disrupted everyday life (1.10 logits). GP attitude (0.86) was perceived to be more important than the available treatments (0.45) or assessment/investigations (0.48). The decision to consult the GP for joint pain was also influenced by other concurrent health problems (0.46) and, to a lesser degree, episodes of more severe, unpredictable pain (0.16).

CONCLUSIONS: Service-related factors appear to be just as influential as the most important need-related determinant of consultation (i.e. pain disruption). Believing the GP would regard joint pain as 'part of the normal ageing process that one just has to accept' is a strong disincentive to seeking help, potentially outweighing other aspects of quality of care. Partial-profile conjoint methods are acceptable to respondents, well-completed, and can address attributes that are unlikely to be accessible in traditional epidemiological designs.

Appendix 1b: Final search terms and databases in systematic review

Medline 1950 to August 2009 search strategy:

Conjoint analysis

Free text words:
1 (conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint)
GP consultation
MESH headings: MH 'Primary Healthcare+', MH 'Referral and Consultation, MM 'Physicians, Family, 'Patient-Centred Care') or (MM 'Continuity of Patient Care, MH 'Decision-making' MH 'Patient Satisfaction+
Free text words: general practitioner, patient preference*

Limit (Human) Total **356**

CINAHL 1982 to August 2009 search strategy:

Conjoint analysis

Free text words:
1 (conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint)
GP consultation
CINAHL headings: MH 'Primary Healthcare+', MH 'Referral and Consultation, MM 'Physicians, Family, 'Patient-Centred Care') or (MM 'Continuity of Patient Care, MH 'Decision-making' MH 'Patient Satisfaction+

Free text words: general practitioner, patient preference*

Limit (Human) Total **147**

Allied & Complementary Medicine (AMED) 1985 to August 2009

Conjoint analysis

Free text words:
1 (conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint)

Total 17

PsycINFO: 1806 to August 2009 Search strategy

Conjoint analysis

Free text words:
1 (conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint)

GP consultation
Thesaurus terms: DE 'Decision-making' OR DE 'Choice Behaviour' OR DE 'Group Decision-making' OR DE 'Management Decision-making', MM Primary Healthcare.
Free text words; TX patient preference*, general practitioner, Patient Satisfaction

Total 284

International Bibliography of social sciences 1951 to August 2009 Search strategy

Conjoint analysis

Free text words:
1 (conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint)

GP consultation
Subjects: (DE 'Health' OR DE 'Child health' OR DE 'Health inequality' OR DE 'Hygiene' OR DE 'Mental health' OR DE 'Public health' OR DE 'Women's health' OR DE 'Illness' OR DE 'Occupational health') and (S20 or S21 or S22 or S23) Search modes - Boolean/Phrase
S23 DE 'Decision-making' OR DE 'Consensus' OR DE 'Conspiracy' OR DE 'Group decision-making' OR DE 'Judgment'. DE 'Primary healthcare'
Free text words: TX patient preference*, general practitioner, patient satisfaction

Total 105

Embase 1974 to august 2009 search strategy:

Conjoint analysis

Free text words:
1 (conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint)
GP consultation
Thesaurus terms: Major descriptor headings: Primary Healthcare, General practitioner, Decision-making, PATIENT CARE, patient satisfaction.

Total 60

Academic Search Premier (1965, to August 2009)

Conjoint analysis

Free text words:
conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint
GP consultation
Subject terms: DE 'PRIMARY healthcare, DE 'PATIENT satisfaction' DE 'PHYSICIANS (General practice)', MH 'Referral and Consultation+, MH 'Decision-making+.
Free text words: general practitioner, patient preference*, patient satisfaction
Total: 261

AGELINE (1975 to August 2009)

Conjoint analysis

Free text words:
1 (conjoint analys* or conjoint N2 analys* or best-worst N2 scal* or Choice-format N2 conjoint-analysis N2 survey* or trade-off N2 analysis or trade-off preference* or revealed preference* or stated preference* or choice-based conjoint or choice-based N2 conjoint or discrete choice* or discrete N2 choice N2 experiment* or discrete N2 choice N2 model* or Adaptive conjoint analysis or Adaptive N2 conjoint N2 analysis or self N2 explicated or Adaptive AND Choice-based AND conjoint or Adaptive Choice-based conjoint)

Total 37

(overall total = 1267)

Appendix 1c: Importance of attributes and levels used in reviewed studies

Attribute	Patient-centred care attributes	Interpersonal and technical care attributes	Process attributes
‡1 (most important attribute)	• **Cheraghi-Sohi et al. 2008 PCC questionnaire:** Physician's knowledge of patient (The doctor has access to your medical notes and knows you well, The doctor has access to your medical notes but does not know you) • **Gerard et al. 2008:**Professional person (Nurse, Doctor-any available, Doctor of choice) • **Rubin et al. 2006:** Choice of doctor (Your choice of doctor, with any available doctor) • **Haas. 2005:** Doctor is trustworthy (Yes, No) • **Hjelmgren et al. 2007:** Patient influence over care received (Large influence, limited influence) • **Fiebig et al. 2009:** Recommended screening interval (1 year, 2 years, 3 years, 5 years)	• **Longo et al. 2006:** Doctor listens (Doctor does not seem to listen, Doctor seems to listen) • **Morgan et al. 2000:** Doctor's manner (Doctor seems to listen, doctor seems to be rushed) • **Vick and Scott 1998:** Being able to talk to the doctor (Doctor does not listen to what you have to say, Doctor listens to what you have to say) • **Scott et al. 2003:** Whether the doctor seems to listen to what you have to say (The doctor seems to listen, The doctor does not seem to listen) • **Scott and Vick 1999:** Being able to talk to the doctor (The Doctor does not seem to listen to what you have to say, Doctor seems to listen to what you have to say) • **Cheraghi-Sohi et al. 2008 Generic questionnaire:** Thoroughness of physical examination (The doctor gives you a thorough examination, The doctor's examination is not very thorough) • **Markham et al. 1999:** Medical skill of the doctor (Good, very good, exceptional)	• **Ryan et al. 1998:** Waiting time between making a non-urgent appointment and seeing a doctor (1 day, 3 days) • **Caldow et al. 2007:** Who you see (a doctor, a practice nurse) • **Turner et al. 2007:** Type of professional consulted (You consult a GP, You consult a nurse)
‡For each of the 15 reviewed studies the attributes have been ranked in order of their relative importance within the individual study. For example, in the Cheraghi-Sohi et al. (2008) study the patient-centred care questionnaire found that the most important attribute was physician's knowledge of patient (as above). This was followed by the patient perspective (see Appendix 1c cont), biopsychosocial perspective (see Appendix 1c cont) shared decision-making (see Appendix 1c cont) and number of days wait for an appointment (see Appendix 1c cont)			

Appendix 1c cont: Importance of attributes and levels used in reviewed studies

Attribute	Patient-centred care attributes	Interpersonal and technical care attributes	Process attributes
2nd most important attribute	• **Longo et al. 2006:** How easily information is understood (Difficult to understand, Easy to understand) • **Cheraghi-Sohi et al. 2008:** **Generic questionnaire:** Physician's knowledge of patient (Doctor has access to your medical notes and knows you well, Doctor has access to your medical notes but does not know you) • **Cheraghi-Sohi et al. 2008: PCC questionnaire:**Patient perspective (The doctor is interested in your own ideas about what is wrong**,** The doctor is not interested in your own ideas about what is wrong) • **Hjelmgren et al. 2007:** Choice for individuals (individual choice of provider-GP or team-, No choice) • **Caldow et al. 2007:** Continuity of health professional (Yes, No) • **Turner et al. 2007:** Relational continuity (Who you do not know, Who you do know) • **Fiebig et al. 2009:** This GP is (Your regular GP for most care, including Pap tests* A GP you do not usually see/have not seen before)	• **Scott and Vick 1999:** Doctor's explanation of information (the doctor's words and explanation are difficult to understand, the doctor's words and explanation are easy to understand) • **Haas. 2005:** Doctor takes notice of what you say (Yes, No)	• **Gerard et al. 2008:**Time of day of appointment (convenient, inconvenient) • **Morgan et al. 2000:**Type of consultation (Home visit by GP, Attendance at an emergency centre run by GP's, attendance at an A and E department, Telephone advice from a GP) • **Rubin et al. 2006:** Time to appointment (Same day, within 48 hours, 4 days, 10 days) • **Vick and Scott 1998:** Information about the treatment of your health problem (The doctor gives you a little information, The doctor gives you a lot of information) • **Markham et al. 1999:** Amount of billing problems (Never, sometimes) • **Ryan et al. 1998:** How long a patient would usually expect to wait in reception (5 min, 15 min, 25 min) • **Scott et al. 2003:** Where your child is seen (Emergency centre run by GPs, your home, a hospital accident and emergency dept)

Appendix 1c cont: Importance of attributes and levels used in reviewed studies

Attribute	Patient-centred care attributes	Interpersonal and technical care attributes	Process attributes
‡3rd most important attribute	• **Longo et al. 2006:** Who chooses treatment (Doctor chooses treatment, You choose, Both choose) • **Cheraghi-Sohi et al. 2008: PCC questionnaire:** Biopsychosocial perspective (The doctor asks about your social and emotional well-being as well as physical symptoms, The doctor asks about your physical symptoms only) • **Haas. 2005:** Doctor recognises your pain/distress (Yes, No) • **Markham et al. 1999:** Who makes care decision (Doctor alone, you and doctor) • **Ryan et al. 1998:** Doctor of choice (any one of the doctors, doctor of choice) • **Scott et al. 2003:** Who your child sees (A GP from your practice/health centre. A GP who doesn't work at your practice/health centre) • **Fiebig et al. 2009:** This GP is Female, Male	• **Cheraghi-Sohi et al. 2008: Generic questionnaire:** Physician's interpersonal manner (Warm and friendly, Formal and business like)	• **Gerard et al. 2008:** Number of days wait for an appointment (Same day, next day, 2 days, 5 days) • **Morgan et al. 2000:** Time you have to wait for consultation with a doctor (30,60, 90, 120 minute intervals) • **Rubin et al. 2006:** Choice of time (Your choice of time, at a specified time) • **Vick and Scott 1998:** Information about your health problem (The doctor gives you a little information, The doctor gives you a lot of information) • **Scott and Vick 1999:** Number of days you wait for the appointment (Less than 2 days, 2 days or more) • **Hjelmgren et al. 2007:** Waiting time for non-emergency visit (7 days, 4 days, 2 days) • **Caldow et al. 2007:** Waiting time to appointment (No waiting time, 2 days, 4 days, 8 days) • **Turner et al. 2007:** Informational continuity (has information about your full medical history, does not have information about your full medical history)

Appendix 1c cont: Importance of attributes and levels used in reviewed studies

Attribute	Patient-centred care attributes	Interpersonal and technical care attributes	Process attributes
‡4th most important attribute	• **Cheraghi-Sohi et al. 2008**: **PCC questionnaire:** Shared decision-making (The doctor involves you in decisions about treatment, The doctor does not involve you in decisions about treatment) • **Haas. 2005:** Doctor treats you with dignity (Yes, No)	• **Caldow et al. 2007:** Likelihood of having illness cured (75%, 80%. 85%)	• **Longo et al. 2006:** Amount of information (Small amount about health, Moderate amount about problem and its treatment, Large amount) • **Gerard et al. 2008:** Length of appointment in minutes (10 mins, 20 mins) • **Morgan et al. 2000:**Ease of getting through to the service (straightforward, complicated) • **Vick and Scott 1998:** Doctor's explanation of information (the doctor's words and explanation are difficult to understand, the doctor's words and explanation are easy to understand) • **Markham et al. 1999:** Time to get a referral (Same day, within 3 days) • **Ryan et al. 1998:** Practice gives patient a health card (no health card offered, health card offered) • **Scott et al. 2003:** Time taken between the telephone call and treatment being received (20 Min, 40 Min, 60 Min, 80 Min) • **Scott and Vick 1999:** Information about your health problem and its treatment (The doctor gives you a little information, The doctor gives you a lot of information) • **Hjelmgren et al. 2007:** Primary care work model (Registration with GP, registration with primary care team) • **Turner et al. 2007:** Access (You are seen on the same day, You wait 2 days for the consultation, You wait 5 days for the consultation, You wait 10 days for the consultation) • **Fiebig et al. 2009:** Doctor's incentive payment (No, Yes)

Appendix 1c cont: Importance of attributes and levels used in reviewed studies

Attribute	Patient-centred care attributes	Interpersonal and technical care attributes	Process attributes
‡5th most important attribute	• **Haas. 2005:** Doctor reassures you (Yes, No) • **Scott and Vick 1999:** Who chooses your treatment (The doctor chooses the treatment for you, The doctor chooses the treatment considering your opinion, You and the doctor make a joint decision, You choose considering doctors opinion) • **Fiebig et al. 2009:** Doctor's recommendation (Not to have a cervical screening test today, Have the standard Pap test today, Have the liquid based cytology Pap test today, Have cervical screening test today but does not recommend one test over other)		• **Longo et al. 2006:** Length of consultation (Less than 10 minutes, More than 10 minutes) • **Cheraghi-Sohi et al. 2008: Generic questionnaire:** Flexibility of appointment times (One appointment offered, Choice of appointment times offered) • **Cheraghi-Sohi et al. 2008: PCC questionnaire:** Number of days wait for an appointment: (Same day, next day, 2 days, 5 days) • **Morgan et al. 2000:** Who you see (doctor you know, doctor you don't know) • **Vick and Scott 1998:** Length of time to see the doctor (10 min, 15 min) • **Markham et al. 1999:** Time in the waiting room (More than 20 mins, less than 20 mins) • **Caldow et al. 2007:** Length of consultation (5 min, 10 min, 20 min, 30 min)

Appendix 1c cont: Importance of attributes and levels used in reviewed studies

Attribute	Patient-centred care attributes	Interpersonal and *technical care* attributes	Process attributes
‡6th most important attribute	• **Vick and Scott 1998:** Who chooses your treatment (The doctor chooses the treatment for you, You both make a joint decision, The doctor lets you choose the treatment)	• **Fiebig et al. 2009:** Time since last cervical screening test (About 1 year, About 2 years, About 3 years, About 5 years)	• **Morgan et al. 2000:** Doctor's shift arrangements (Doctor is awake and is only on duty until the morning, On call doctor is woken up and will be working the next day) • **Haas. 2005:** Doctor gives you information (if you ask for it, whether you ask for it or not, Only where you can get information, No)
‡7 (least important attribute)	• **Haas. 2005:** Doctor accepts your decisions about your health (Yes, Yes but also gives advice or opinion, No but tells you about his/her decisions, No)		
‡For each of the 15 reviewed studies the attributes have been ranked in order of their relative importance within their individual study. For example, in the Cheraghi-Sohi et al. (2008) study the patient-centred care questionnaire found that the most important attribute was physician's knowledge of patient (see Appendix 1c cont). This was followed by the patient perspective (see Appendix 1c cont), biopsychosocial perspective (see Appendix 1c cont) shared decision-making (see Appendix 1c cont) and number of days wait for an appointment (see Appendix 1c cont)			

Appendix 2: Results of inter-rater agreement

Included studies	Item part	ISPOR checklist item 1	ISPOR checklist item 2	ISPOR checklist item 3	ISPOR checklist item 4	ISPOR checklist item 5	ISPOR checklist item 6	ISPOR checklist item 7	ISPOR checklist item 8	ISPOR checklist item 9	ISPOR checklist item 10
1. Caldow et al. (2006)	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	Y	Y	Y	U	Y	Y	Y	Y	Y	Y
	3	Y	Y	Y	Y	N	Y	Y	Y	Y	Y
2.Cheraghi-Sohi et al. (2008)	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
	3	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
3. Fiebig (2009)	1	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
	2	Y	Y	U	U	Y	Y	Y	U	Y	N
	3	Y	Y	U	Y	N	N	Y	Y	N	Y
4. Gerard et al.(2008)	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	3	Y	Y	Y	N	N	Y	U	Y	N	Y
5. Haas. (2005)	1	Y	Y	Y	N	Y	Y	Y	Y	N	Y
	2	Y	U	Y	N	Y	Y	Y	N	Y	N
	3	Y	Y	U	U	N	Y	Y	Y	Y	U
6.Hjelmgren et al.(2007)	1	Y	Y	Y	Y	Y	Y	U	Y	N	Y
	2	Y	Y	Y	N	Y	U	Y	N	Y	N
	3	Y	Y	Y	N	N	Y	U	Y	U	Y
7. Longo et al.(2006)	1	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
	2	Y	Y	Y	Y	Y	U	Y	Y	Y	N
	3	Y	Y	Y	Y	N	Y	Y	Y	U	Y
8. Markham et al.(1999)	1	N	U	Y	U	Y	N	N	Y	N	Y
	2	Y	U	U	N	N	N	Y	Y	N	U
	3	Y	U	Y	N	N	Y	U	U	N	Y

Appendix 2 cont: Results of inter-rater agreement

Included studies	Item Part	ISPOR checklist item 1	ISPOR checklist item 2	ISPOR checklist item 3	ISPOR checklist item 4	ISPOR checklist item 5	ISPOR checklist item 6	ISPOR checklist item 7	ISPOR checklist item 8	ISPOR checklist item 9	ISPOR checklist item 10
9. Morgan et al.(2000)	1	U	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	Y	U	Y	U	Y	Y	Y	N	Y	N
	3	Y	Y	Y	N	N	Y	U	U	Y	Y
10. Rubin et al.(2006)	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	Y	U	U	N	Y	Y	Y	Y	Y	Y
	3	Y	U	Y	U	N	Y	Y	Y	Y	Y
11. Ryan et al.(1998)	1	Y	Y	Y	Y	Y	U	Y	Y	Y	Y
	2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	3	Y	Y	Y	N	N	Y	Y	U	Y	Y
12. Scott and Vick (1999)	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	Y	Y	Y	U	Y	Y	U	Y	Y	Y
	3	Y	Y	Y	N	N	N	Y	Y	U	Y
13. Scott et al.(2003)	1	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	2	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
	3	Y	Y	Y	U	N	Y	Y	Y	Y	Y
14. Turner et al.(2007)	1	Y	Y	Y	Y	Y	Y	Y	Y	N	Y
	2	Y	Y	Y	N	Y	Y	Y	N	Y	Y
	3	Y	Y	Y	U	N	Y	U	Y	Y	Y
15. Vick and Scott (1998)	1	Y	Y	Y	Y	Y	Y	N	Y	Y	Y
	2	Y	Y	Y	Y	Y	Y	U	Y	Y	Y
	3	U	Y	Y	N	N	Y	Y	N	Y	Y

Appendix 3a: Search strategy for preliminary test 1
Medline 1950 to June 2010 search strategy:
1. Osteoarthritis
Free text words:
Osteoarthritis, musculoskeletal, joint pain.
2. Help-seeking behaviour
Abstract Only:
Self regulatory model, Andersen model, predisposing factors, general practitioner, general practitioner consultation, help-seeking, consultation, patient decision-making, service utilisation, perceived need, care seeking.
Limit (Middle aged: 45-64 years)
Limit (Aged: 65+ years)
Limit (Subject Major: Primary Healthcare)
Limit (Subject Major: Musculoskeletal Diseases epidemiology)
Limit (Subject Major: Osteoarthritis)
Limit (Subject Major: Referral and Consultation)
Total: 170

CINAHL 1982 to June 2010 search strategy:
1. Osteoarthritis
Free text words:
Osteoarthritis, musculoskeletal, joint pain.
2. Help-seeking behaviour
Abstract Only:
Self regulatory model, Andersen model, predisposing factors, general practitioner, general practitioner consultation, help-seeking, consultation, patient decision-making, service utilisation, perceived need, care seeking.
Limit (Middle aged: 45-64 years)
Limit (Aged: 65+ years)
Limit (Subject Major: Primary Healthcare)
Limit (Subject Major: Musculoskeletal Diseases epidemiology)
Limit (Subject Major: Osteoarthritis)
Limit (Subject Major: Referral and Consultation)
Total: 67

Allied & Complementary Medicine (AMED) 1985 to June 2010
1. Osteoarthritis
Free text words:
Osteoarthritis, musculoskeletal, joint pain.
2. Help-seeking behaviour
Boolean/Phrase mode:
Self regulatory model, Andersen model, predisposing factors, general practitioner, general practitioner consultation, help-seeking, consultation, patient decision-making, service utilisation, perceived need, care seeking.
Limit (PRIMARY HEALTHCARE) or
Limit (OSTEOARTHRITIS)
Total: **15**

PsycINFO: 1806 to June 2010 Search strategy
1. Osteoarthritis
Free text words:
Osteoarthritis, musculoskeletal, joint pain.
2. Help-seeking behaviour
Boolean/Phrase mode:
Self regulatory model, Andersen model, predisposing factors, general practitioner, general practitioner consultation, help-seeking, consultation, patient decision-making, service utilisation, perceived need, care seeking.
Limit (Middle age: 40-64 years)
Limit (Aged: 65 years and older)
Total: 71

British Nursing Index 1994 to June 2010 search strategy:
1. Osteoarthritis
Free text words:
Osteoarthritis, musculoskeletal, joint pain.
2. Help-seeking behaviour
Boolean/Phrase mode:
Self regulatory model, Andersen model, predisposing factors, general practitioner, general practitioner consultation, help-seeking, consultation, patient decision-making, service utilisation, perceived need, care seeking.
Total: 7

SPORTDiscus with Full Text 1985 to June 2010
1. Osteoarthritis
Free text words:
Osteoarthritis, musculoskeletal, joint pain.
2. Help-seeking behaviour
Boolean/Phrase mode:
Self regulatory model, Andersen model, predisposing factors, general practitioner, general practitioner consultation, help-seeking, consultation, patient decision-making, service utilisation, perceived need, care seeking.
Limit (PRIMARY care (Medicine) or
Limit (PATIENTS)
Total: 24

Academic search premier 1985 to June 2010
1. Osteoarthritis
Free text words:
Osteoarthritis, musculoskeletal, joint pain.
2. Help-seeking behaviour
Boolean/Phrase mode:
Self regulatory model, Andersen model, predisposing factors, general practitioner, general practitioner consultation, help-seeking, consultation, patient decision-making, service utilisation, perceived need, care seeking.
Limit (MEDICAL consultation)
Total: 6

Appendix 3b: Cognitive interview script; introduction/ending

Introduction

PLEASE ENSURE YOU COMMUNICATE THE HIGHLIGHTED KEY WORDS TO RESPONDENT

- **Hello and Thank you** for agreeing to complete the questionnaire and coming along to help us today.

- **My name is** …….and I am **working on behalf of Keele University**. I'll just take a few moments to go over the reasons why we are conducting the study: ***We are interested in patients decision-making around consulting the doctor for their joint pain*** *and we have developed a questionnaire to be able to do this.* ***But before we send this questionnaire by post*** *to patients across North Staffordshire it is important to find out if the questions we are asking are* ***clear and easy to understand.***

- As this is a **postal questionnaire** it is important to us that is it is easy to understand.

- **This is where we need your help.** *It is only by asking people like you who have foot problems and by getting them to give your* ***advice on the questions****, that we can be sure we are asking the right questions and that the questions work in the way that they are intended.*

- Our main purpose today is to get your **advice** about this questionnaire as it has not been used before, so this is a test of the questionnaire and **NOT OF YOU**. **There are NO RIGHT or WRONG answers.**

- I would like to **record** this interview so I can pay attention to you without having to make notes, is this ok with you?

- The information you provide on the questionnaire and verbally will be treated in strict confidence. The researchers may use data from the questionnaire and quotes from this interview when reporting the findings but these will not identify you. This tape recording is entirely **confidential** and will be kept at keele university holdings safely and securely.

- As you are completing the questionnaire what I would like you to do is to fill in the questionnaire and at the same time **tell me what you are thinking**. So just say out loud what you are thinking while you read the questions and put in your answers. I would like you to speak aloud any thoughts that you may have about the questions, both positive and negative.

- **WINDOWS EXERCISE**; In order to get you into the way of thinking that we are looking for would you like to go through a mini training exercise with me now? '***I want you to close your eyes and try to visualise the place where you live, and think about how many windows there are in that***

place. As you count up the windows, tell me what you are seeing and thinking about'

- *Thank you for doing that. How did that feel?* So please try to tell me what you are thinking as you approach and complete the questions.

- In particular, please comment if the responses options available to you do not allow you to answer the question in the way that you would like to, for example (show example). Please make it clear if you do not understand the question, or indeed if you feel that something is missing. All feedback will be greatly appreciated.

- There are **no right or wrong answers** to the questions - what is important is that you answer the questions as honestly as you can.

- If you normally read aloud when completing questionnaires like this then please do this here today. Don't read aloud the ones that you normally would not. Basically, do what you would normally do when filling in questionnaires that you receive through the post.

- Do you feel happy with what you have to do?

- Do you have any questions?

PROCEED INTO INTERVIEW SECTION A
Tips for Researcher;

- At the end of each section reassure and reward respondent.

- Pay close attention and if respondents look confused use general probes to elicit further information, If respondents look tired, suggest a break.

- Do NOT interrupt the thinking aloud process with additional questions; It is your ability as a researcher to allow their natural flow to take place. Bite your lip!

- Try not to treat the respondent as if they are expected to be an expert on questionnaire layout.

Ending

- Ok, thank you very much for going through the questionnaire with us today. Your feedback is really importance to us and your views will be processed and used when we finalise the questionnaire.

- Do you have any questions or is there anything you would like to add? (for example, the overall length of the questionnaire)

- **REITERATE**; This tape recording is entirely **confidential** and will be kept at keele university holdings safely and securely.

- How was that for you?

Appendix 3c: Cognitive Interview script; pre-planned probes for preliminary test 1

Introduction

Instead of explaining the technical principles of the DCE relay the exercise (i.e. choosing a package holiday scenario) to respondents to orient them to the task.

Relevance of question: rating exercise

- how relevant or irrelevant did you find those attributes?
- are there any other attributes you would like to add to this section?

Response : DCE

- how easy or difficult did you find answering this/that question? why was that?
- how did you feel about answering this/these questions?
- did you find any of the questions too personal or too intrusive?

Length: DCE

- How did you find the overall length?
- Were 5 choice sets manageable?
- Could you have managed any more?

Appendix 3d: Cognitive Interview script; general probes

Think aloud **GENERAL** Probes (to be used in a non-directional manner)

- Remember to tell me what you're thinking.
- How did you arrive at that answer?
- What went on in your mind when you were asked that question?
- Could you tell me more about your thinking on that?
- How do you mean?

Rewarding:

- That's exactly the kind of information we need. Thank you for that.
- Thank you, that is what we want to hear.

Comprehension 1; paraphrasing:

- Can you repeat that question in your own words?
- What would you say that question was asking of you?

Comprehension 2:

- In your words, what is…..?
- What does ……..mean to you?

Response

- How easy or difficult did you find answering this/that question? why was that?
- How did you feel about answering this/these questions
- Did you find any of the questions too personal or too intrusive?

Visual layout

- I noticed you were looking here. what were you thinking?
- I noticed you were looking here. what led you to do that?
- How easy or difficult did you find it to fill out this questionnaire this questionnaire?

Recall/judgment

- What brought that to mind?
- How did you work that out?
- How did you remember that?
- Did you try and count each time, or did you make an estimate?
- What time period were you thinking of? (from when to when?)

- What did you think of as you tried to remember (reference time period)?
  Relevance of question

- How relevant or relevant did you find that question/section?

- How important or unimportant did you find that question/section?

Appendix 3e: CBC/DCE used in Preliminary test 1

1. Rating exercise

I want you to imagine that you have developed joint pain in various parts of your body. Occasionally it is more than just minor pain. The problem does not appear to be getting consistently worse but it has not responded to exercise or advice recommended by the chemist. You are deciding whether to seek a medical opinion from a GP. Please rate the factors below in order of importance in your decision to consult from 1 (least important) to 10 (most important).

If there are any factors that you consider to be important that are missing please add them into the boxes provided at the bottom of this table and rank them using the same scale.

Factor	**Rating (1 - 10)**
1. Severity of Pain (mild, moderate, severe)	
2. Whether you have other illnesses (hypertension, diabetes)	
3. How likely it is that you will be helped/cured. (pain reduced, physical functioning/emotional functioning)	
4. The doctor has access to your medical notes and knows you well/not very well	
5. Thoroughness of physical examination	
6. The doctor's interest in your own ideas about what is wrong	
7. Time it takes to get an appointment.	
8. ADD in here	
9. ADD in here	
10. ADD in here	

2. Discrete choice exercise

There are 5 choice sets in total.

Scenario 1

I want you to imagine that you have developed joint pain in various parts of your body. Occasionally it is more than just minor pain. The problem does not appear to be getting consistently worse but it has not responded to exercise or advice recommended by the chemist. You are deciding whether to seek a medical opinion from a GP. Please choose between the options below (A or B) as to which best represents your decision to consult the GP.

.

Choice set 1. If you were offered options A and B below which one would you choose?

Option A

- My joint pain is mild
- I have no other illnesses
- The doctor can help to Increase my physical functioning
- The doctor has access to my medical notes but does not know me

Option B

- My joint pain is moderate
- I have one other illnesses (eg, hypertension/diabetes)
- The doctor can help to Increase my physical functioning
- The doctor has access to my medical notes and knows me a little

Choice

Choice Set 2. If you were offered options A and B below which one would you choose?

Option A	Option B
• My joint pain is mild • I have no other illnesses. • The doctor can help to reduce my Pain (severity, character and frequency) • The doctor has access to your medical notes and knows you well	• My joint pain is mild • I have one or more other illnesses (hypertension, diabetes) • The doctor can help to increase my emotional functioning • The doctor has access to your medical notes but does not knows me

Choice

Choice Set 3. If you were offered options A and B below which one would you choose?

Option A	Option B
• My joint pain is Moderate • I have one other illness (hypertension, diabetes) • The doctor can help to Increase my emotional functioning • The doctor has access to my medical notes and knows me well	• My joint pain is severe • I have one or more other illnesses (hypertension, diabetes) • The doctor can help to Increase my physical functioning • The doctor has access to my medical notes and knows me well

Choice

Choice Set 4. If you were offered options A and B below which one would you choose?

Option A	Option B
• My joint pain is severe • I have one other illness (hypertension, diabetes) • The doctor can help to reduce my pain (severity, character and frequency) • The doctor has access to my medical notes but does not know me	• My joint pain is moderate • I have one or more other illnesses (hypertension, diabetes) • The doctor can help to reduce my pain (severity, character and frequency) • The doctor has access to my medical notes and knows me a little

Choice

Choice Set 5. If you were offered options A and B below which one would you choose?

Option A	Option B
• My joint pain is severe • I have no other illnesses • The doctor can help to Increase my emotional functioning • The doctor has access to my medical notes and knows me a little	• My joint pain is mild • I have no other illnesses. • The doctor can help to reduce my Pain (severity, character and frequency) • The doctor has access to your medical notes and knows you well

Choice

4. Demographic characteristics of patient sample;

- **Sex**
- **Age**
- **Ethnicity**
- **Number of dependent children in household**
- **Have OA/Joint related conditions**
- **Have any other long-term illnesses**
- **Overall health is 'good' or 'very good,'**
- **Mental health is 'good' or 'very good,'**
- **Number of visits to physician in past 12 months**

Do you have any additional comments?.....

Thank you for your help with completing this questionnaire.

Appendix 3f: Information sheet sent to respondents prior to preliminary test 1

KEELE UNIVERSITY

arc Committed to curing arthritis

NHS National Institute for Health Research

PRIMARY CARE SCIENCES
ARTHRITIS RESEARCH CAMPAIGN NATIONAL PRIMARY CARE CENTRE

Address

Date

Dear respondent

New study on how people make choices in healthcare

Thank you for your recent help in the foot study questionnaire. We are writing up the notes from these meetings and will send them to you soon. In the meantime we wondered if you wanted to help us with this new project.
The new study aims to find out about the choices that people make when deciding whether or not to raise their joint pain problem with the general practitioner. We want to ask people like you about two things:

1) To tell us how important different things are to the decision to consult the GP. Examples would be how bad the pain is, how long it's been going on, expectations about what your doctor can do about it.
2) To compare a series of different scenarios and choose which one you think would be more likely to result in you consulting the GP

If you agree to help and give us feedback then it will involve you coming to the Centre here at Keele and filling out a questionnaire. The questionnaire should take you about 45 minutes to complete and then we will ask you about how it went. We will ask you how easy or hard it was to fill in, and how easy or hard it was to make choices about different scenarios. As this is a new type of survey your advice will be extremely valuable and we will use your feedback to improve the methods we use in a larger study that we are planning to undertake next year.

We would be very grateful if you would let us know if this is something you would be interested in helping us with. If you would like to know more about this study, please contact Carol Rhodes at Keele University on c.a.rhodes@cphc.keele.ac.uk or telephone on 01782 734834 Tuesday, Wednesday, or Thursday. Or please fill in the attached form and return it in the prepaid envelope provided.
Thank you very much for reading this letter

Yours sincerely,

Researcher

T: +44(0)1782 733905 F: +44(0)1782 733911
E: primary_care_sciences@cphc.keele.ac.uk W: www.keele.ac.uk/research/pchs/pcmrc/

Keele University, Staffordshire, ST5 5BG, United Kingdom
T: +44(0)1782 732000 or 621111 W: www.keele.ac.uk

Appendix 4a: ACBC Task

Sawtooth Software SSI Web Demo

Please select the situation that will most likely lead you to consult your GP regarding your joint pain. For each feature please select your preferred answer.

Feature	Select Feature
Pain compared to usual	◉ The pain I am experiencing is less than usual ○ The pain I am experiencing is the same as usual ○ The pain I am experiencing is more than usual
Pain relief	◉ 15% pain relief (minimal) ○ 30% pain relief (moderate) ○ 50% pain relief (substantial)
GP you can see	○ You see the same GP ○ You see a different GP
GP Manner	○ The GP is friendly and informal ○ The GP is businesslike
Information on joint pain given	○ The GP provides you with no information about your joint pain ○ The GP provides you with a little information about your joint pain ○ The GP provides you with a lot of information about your joint pain
GP offers you	○ GP offers you painkillers only ○ GP offers you painkillers and physiotherapy

Sawtooth Software SSI Web Demo

Here are a few situations that might encourage you to consult your GP regarding your joint pain. For each one indicate whether it is a possibility or not.

Pain compared to usual	The pain I am experiencing is the same as usual	The pain I am experiencing is less than usual	The pain I am experiencing is more than usual	The pain I am experiencing is less than usual
Pain relief	30% pain relief (moderate)	30% pain relief (moderate)	15% pain relief (minimal)	50% pain relief (substantial)
GP you can see	You see the same GP	You see a different GP	You see the same GP	You see a different GP
GP Manner	The GP is friendly and informal	The GP is businesslike	The GP is friendly and informal	The GP is friendly and informal
Information on joint pain given	The GP provides you with no information about your joint pain	The GP provides you with a lot of information about your joint pain	The GP provides you with a little information about your joint pain	The GP provides you with no information about your joint pain
GP offers you	GP offers you painkillers and physiotherapy	GP offers you painkillers only	GP offers you painkillers only	GP offers you painkillers and physiotherapy
	○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me	○ A possibility ○ Won't work for me

Sawtooth Software SSI Web Demo

Among these three, which is the best option? (I've greyed out any features that are the same, so you can just focus on the differences.)

Pain compared to usual	The pain I am experiencing is more than usual	The pain I am experiencing is the same as usual	The pain I am experiencing is the same as usual
Pain relief	30% pain relief (moderate)	15% pain relief (minimal)	30% pain relief (moderate)
GP you can see	You see the same GP	You see the same GP	You see the same GP
GP Manner	The GP is friendly and informal	The GP is businesslike	The GP is friendly and informal
Information on joint pain given	The GP provides you with no information about your joint pain	The GP provides you with a lot of information about your joint pain	The GP provides you with a little information about your joint pain
GP offers you	GP offers you painkillers only	GP offers you painkillers only	GP offers you painkillers and physiotherapy

Next

Appendix 4b: Priority worksheet for RUG

1. 6 factors used in computer questionnaire used today

Attribute	Attribute levels
Pain compared to usual	(●) The pain I am experiencing is less than usual ( ) The pain I am experiencing is the same as usual ( ) The pain I am experiencing is more than usual
Pain relief	(●) 15% pain relief (minimal) ( ) 30% pain relief (moderate) ( ) 50% pain relief (substantial)
GP you can see	( ) You see the same GP ( ) You see a different GP
GP Manner	( ) The GP is friendly and informal ( ) The GP is businesslike
Information on joint pain given	( ) The GP provides you with no information about your joint pain ( ) The GP provides you with a little information about your joint pain ( ) The GP provides you with a lot of information about your joint pain
GP offers you	( ) GP offers you painkillers only ( ) GP offers you painkillers and physiotherapy

2. **Are there any other factors which are missing that are relevant that you would like to see included? (15 minutes)**

3. **So, we have a new list of factors (above). I want you to imagine you are the researcher and you have to choose only 6.**

   1.
   2.
   3.
   4.
   5.
   6.

Appendix 4c: Pre-planned probes used in preliminary test 2

Specific objectives	Action and probes	Required materials
○ To uncover issues of 'comprehension, recall, judgment and response' in the completion of ACBC exercise using Think aloud.	○ **Volunteer 1 Think aloud and complete ACBC exercise** ○ **Inform users that there will be a group discussion at the end of each slide.** ○ **Slide 2: How comprehensive are they? How relevant are they?** ○ **At the end of each slide 'what did you think of that? How easy or difficult was that to do?'** ○ **Probe around the levels of each attribute, i.e. were they understandable?** ○ **How easy or difficult was it to trade against the pain?**	Computer and projector X 9 pens and paper Dictaphone
○ To uncover issues of 'relevance' and 'comprehension' (reasons for consultation) that arose ○ To ask what other factors might be important in addition to 6 attributes used. ○ Play role of researcher and decide upon 6 attributes as most important.	○ **Give out worksheet** ○ **Look at the 6 factors used, were they relevant? Are there any missing?** ○ **Which other factors may influence your decision to consult the GP for joint pain?** ○ **Use probe sheet (for other factors)** ○ **Pretend you are a researcher and of this list can you decide individually on which 6 may be the most important?**	X 9 pens and paper Dictaphone Worksheet
○ To discuss issues of 'method' (ACBC) that arose in the completion of ACBC exercise	○ **Open ended question 'what did you think of that. How easy or difficult was that on the computer?'** ○ **Did it reflect your thoughts? Was it listening to you? Was it relevant and engaging?** ○ **How was it using the computer?** ○ **Did you need any computer skills to fill it in?** ○ **Was the layout clear?**	x9 pens and paper Dictaphone Stopwatch

Appendix 4d: Information sheet sent to respondents prior to preliminary test 3

KEELE UNIVERSITY

arc
Committed to curing arthritis

NHS
National Institute for Health Research

PRIMARY CARE SCIENCES
ARTHRITIS RESEARCH CAMPAIGN NATIONAL PRIMARY CARE CENTRE

Address

Date

Dear respondent

Choice questionnaire on how people make choices when deciding to consult the GP for joint pain

We are writing to you as a member of the Research Users' Group at the arc National Primary Care Research Centre. We wondered if you wanted to help us in our final stage of developing our new choice-based experiment.

If you remember, the new study aims to find out about the choices that people make when deciding whether or not to raise their joint pain problem with the general practitioner. We are looking for one volunteer to compare a series of different situations and choose which one they think would be more likely to lead you to consult the GP for your joint pain. One volunteer will be asked to 'think aloud' during this process, and the rest of the group will be invited to join in a group discussion at the end of each page.

We want to ask you about three things:

1) To tell us how effective sections A,B and C (part 1) are in explaining the task and preparing participants for making choices. Little time will be spent on this in the session so if you would look over this section and we will ask for some feedback at the start of the session.
2) To tell us how effective Section C is overall, i.e. the choice questionnaire. We will be focussing mainly on how the factor 'other health problems' work in this design. We have categorised other health problems into 'major' and 'minor' and want to discuss if these are appropriately categorised.
3) To tell us how effective the layout, length and other elements of the design are when making choices.

The final section 'about you' is not going to be a focus for the meeting, as this a well used questionnaire within the centre.

If you agree to help and give us feedback then it will involve you coming to the Centre here at Keele and partaking in a group discussion after a 'think aloud' has been done by one volunteer. The session will last for about 60 minutes and we will ask you about how it

went. We will ask you how easy or hard it was to fill in, and how easy or hard it was to make choices about different situations. As this is a new type of survey your advice will be extremely valuable and we will use your feedback to improve the methods we use in conducting this postal study.

We would be very grateful if you would let us know if this is something you would be interested in helping us with. If you would like to know more about this study, please contact Carol Rhodes at Keele University on 01782 73

Thank you very much for your help with this research project.

Yours sincerely,

Researcher

T: +44(0)1782 733905 F: +44(0)1782 733911
E: primary_care_sciences@cphc.keele.ac.uk W: www.keele.ac.uk/research/pchs/pcmrc/

Keele University, Staffordshire, ST5 5BG, United Kingdom
T: +44(0)1782 732000 or 621111 W: www.keele.ac.uk

Appendix 5a: CBC used in preliminary test 3.

KEELE UNIVERSITY

Arthritis Research UK Primary Care Centre
working with XXXX GP practice

A QUESTIONNAIRE ABOUT THE DECISION TO CONSULT THE GENERAL PRACTITIONER FOR JOINT PAIN

Version 1; August 2010
Research Ethics Committee reference: XX/XXXX

Section A. Introduction

Thank you for helping with this study. The aim of this questionnaire is to understand what leads people with joint pain to consult a general practitioner about it. Even if you don't see a GP very often, we would be very grateful if you would take part as it is important that we hear from a wide variety of people.

In this questionnaire there are three sections. Firstly, we would like you to consider the importance of some factors that may influence your decision to consult the GP for your joint pain. The factors are things that people with joint problems have told us they might consider when deciding whether to seek medical attention for a joint problem. These things include the symptoms you are experiencing at the time and the amount of disruption it is causing to your everyday life.

In the second section we would like you to complete 16 choice tasks, which are made up of two imaginary situations. These imaginary situations are made up of the factors that you have already considered. From these two imaginary situations we would like you to choose which one would more likely lead you to consult the general practitioner.

In the third section we would like you to complete some questions about how you cope, how you think about your joint pain and there are also some general health questions.

It is important to remember that these situations are hypothetical. We ask that you try to 'put yourself in each situation', and choose the one that would more likely lead you to consult the general practitioner.

There are no right or wrong answers. We are interested in your own personal preferences based on how you might feel in each situation.

INSTRUCTIONS FOR THIS QUESTIONNAIRE

Please answer all of the questions.

The questions can be answered by putting a cross in a box lik☒is:

or circling a number like this: 3  5 6

Please write in BLOCK CAPITALS where appropriate

When you have finished please check that you have answered all of the questions and then return the questionnaire in the envelope enclosed. You do not need a stamp. Please return the questionnaire by August 31st 2010.

The answers you give in the questionnaire will be treated in the strictest confidence.

Whether you take part in this research or not, your right to use health services at your practice or elsewhere will not be affected.

If you have any more questions please contact our Study Co-ordinator, Nica, on 01782 4829.

Thank you again for your help with this research study.

Section B. Factors that may influence your decision to consult the GP for your joint pain

We asked people with joint problems which things might influence their decision to go and see the doctor about it. Below is a list of some of the things they came up with. Please read this list and then move on to question 1 on the next page.

a.	The type of joint pain you are experiencing at the time Is it the 'usual aches and pains' or am I getting a different kind of pain, for example short episodes of unpredictable pain?
b.	The severity of the joint pain you are experiencing How bad is the pain? Some people report that their pain fluctuates, they have 'good' and 'bad' days.
c.	The level of disruption that the joint problem is causing to your everyday life How much is the joint problem interfering with my usual activities? Joint problems can disrupt many aspects of everyday living including sleep, daily activities (getting about, looking after the house), family responsibilities, social and leisure activities, and relationships.
d.	Other health problems that may be bothering you at the time Have I got any other health problems at the moment? These could range from short-lasting minor complaints (e.g. a bad cold, headache) to long-term major health conditions (e.g. Diabetes, heart problems).
e.	The sorts of treatments the GP is likely to offer you Is the GP going to offer painkillers? Lifestyle advice? An injection? Complementary therapies? Perhaps there is a new treatment available?
f.	The sorts of referrals and further investigations the GP is likely to offer you Will the doctor consider referring me to a physiotherapist or a hospital specialist? Will they order an X ray?
g.	The GP's general manner Is the GP going to take the time to listen carefully or are they going to be in a rush?
h.	The GP's attitude to joint problems Am I going to be told that it's just 'part of the normal ageing process'?

Please consider the factors below (from a – h) and circle a number from 1 (least important) to 10 (most important) to indicate how important they are in your decision to consult.
(Please circle a number for EVERY factor from a - h)

		Not at all important									Extremely important	
a.	The type of joint pain you are experiencing at the time	0	1	2	3	4	5	6	7	8	9	10
b.	The severity of the joint pain you are experiencing	0	1	2	3	4	5	6	7	8	9	10
c.	The level of disruption that the joint problem is causing to your everyday life	0	1	2	3	4	5	6	7	8	9	10
d.	Other health problems that may be bothering you at the time	0	1	2	3	4	5	6	7	8	9	10
e.	The sorts of treatments the GP is likely to offer you	0	1	2	3	4	5	6	7	8	9	10
f.	The sorts of referrals and further investigations the GP is likely to offer you	0	1	2	3	4	5	6	7	8	9	10
g.	The GP's manner	0	1	2	3	4	5	6	7	8	9	10
h.	The GP's attitude to joint problems	0	1	2	3	4	5	6	7	8	9	10

Section C. Choice questionnaire

Part 1– Choice tasks

For each of the following choice tasks numbered 1-16 in this following section, please carefully read the two imaginary situations (A and B) and chose which ONE (A or B) would MORE LIKELY LEAD YOU TO CONSULT A GP IN YOUR PRACTICE FOR YOUR JOINT PAIN by ticking the relevant box beneath.

Please choose one of the OVERALL situations (A or B) even though neither may seem ideal.

At times some of these situations may overlap with each other and may seem to be repetitive for you, but none of these choices are exactly the same and it is important to us that you consider each one separately.

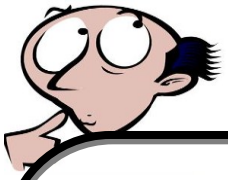
Please note also that when a factor is greyed out in both situations A and B this means that they are the same across both situations. This may help you to make a choice about which situation is most relevant for you.

After choice 8 there is a suggested break period, which will allow you to take a break before continuing to fill out this choice questionnaire. We are aware it is quite a complex task and we do not want to over burden you.

Part 2– How to complete the choice questionnaire:

On the following two pages is an example of how a person with joint pain called Joe completed the choice task. Please use this to guide you if you get stuck at any point in the questionnaire.

Imagine that your joint problem and its impact on your everyday life is changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**



Imaginary Situation A

- **IF** I was experiencing **a dull aching pain, which was there most of the time**
- **IF** the pain was disrupting **most** of my everyday life
- **IF** the GP would offer me **painkillers**
- **IF** the GP would regard my joint pain **as part of the normal ageing process that one just has to accept**

Step 1

Step 2

Step 3

Step 4

Imaginary Situation B

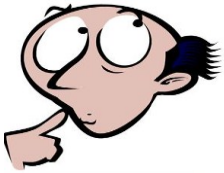
- **IF** I was experiencing **short episodes of an unpredictable pain**
- **IF** the pain was **not** disrupting my everyday life
- **IF** the GP would offer me a **promising new**
- **IF** the GP would regard my joint pain **as part of the normal ageing process that**

I would be more likely to consult the doctor about the joint problem if I was in…

Situation A ☐

Step 5

Situation B ☐



STEP 1

Joe compares the first two factors in situation A and B and decides **short episodes of an unpredictable pain** would more likely lead him to consult the GP

STEP 2

Joe compares the second two factors in situation A and B and decides **If the pain is disrupting most of my everyday living** it would more likely lead him to consult the GP

STEP 3

Next, Joe compares the third two factors in situation A and B and decides 'IF the GP will offer me **a promising 'new' treatment** he would more likely consult the GP

STEP 4

Joe moves on to the fourth two factors in situation A and B and can see that these factors are both the same in situation A and B (this is why they are greyed out). Therefore the fourth two factors are cancelled out in his decision-making process as they are the same

STEP 5:

Joe now has to decide which combination of the factors within both situation A & B overall is the most important in his decision. Joe decides the pain he was in is less important to him than the **disruption to everyday life** and if a '**promising new treatment'** was available, He then has to decide between these two factors. Joe decides that out of these two factors a '**promising new treatment'** would most likely lead him to consult the GP for his joint pain. He therefore chooses situation B.

People have different ways of filling out the choice questionnaire. Some people compare the factors between situations and others will look at the situations as a whole and try to compare the whole situation against the other one. Others simply chose the factors most important to them and overlook less important factors. It is entirely up to you how you make your choices.

Choice Set 1

Imagine that your joint problem and its impact on your everyday life is changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- **IF** I was experiencing **a dull aching pain, which was there most of the time**
- **IF** the pain I was experiencing was **severe**
- **IF** the pain was **not** disrupting my everyday life
- **IF** I was experiencing **no other** physical health problems

Imaginary Situation B

- **IF** I was experiencing **short episodes of an unpredictable pain**
- **IF** the pain I was experiencing was **moderate**
- **IF** the pain was disrupting **some** of my everyday life
- **IF** I was experiencing **other minor** physical health problems

I would be more likely to consult the doctor about the joint problem if I was in…

Situation A ☐ Situation B ☐

Choice Set 2

Imagine that your joint problem and its impact on your everyday life is changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- **IF** I was experiencing a **dull aching pain, which was there most of the time**
- **IF** the pain was **not** disrupting of my everyday life
- **IF** the GP would offer me a **referral to a hospital specialist** (rheumatologist, orthopaedic doctor, pain clinic etc)
- **IF** the GP would **seem to be rushed**

Imaginary Situation B

- **IF** I was experiencing **short episodes of an unpredictable pain**
- **IF** the pain was **not** disrupting my everyday life
- **IF** the GP would offer me a **referral for an X-ray**
- **IF** the GP would **seem to be rushed**

I would be more likely to consult the doctor about the joint problem if I was in…

Situation A ☐

Situation B ☐

Choice Set 3

Imagine that your joint problem and its impact on your everyday life is changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- **IF** the pain I was experiencing was **moderate**
- **IF** I was experiencing **other minor** physical health problems
- **IF** the GP would **take time to listen**
- **IF** the GP would regard my joint pain **as part of the normal ageing process that one just has to accept**

Imaginary Situation B

- **IF** the pain I was experiencing was **mild**
- **IF** I was experiencing **other major** physical health problems
- **IF** the GP would **take time to listen**
- **IF** the GP would regard my joint pain **as a legitimate health problem that requires treatment**

I would be more likely to consult the doctor about the joint problem if I was in…

Situation A ☐

Situation B ☐

Choice Set 4

Imagine that your joint problem and its impact on your everyday life is changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- **IF** the pain I was experiencing was **moderate**
- **IF** the GP would offer me a **referral to a physiotherapist**
- **IF** the GP would **seem to be rushed**
- **IF** the GP would regard my joint pain **as a legitimate health problem that requires treatment**

Imaginary Situation B

- **IF** the pain I was experiencing was **mild**
- **IF** the GP would offer me a **referral to a hospital specialist** (rheumatologist, orthopaedic doctor, pain clinic etc)
- **IF** the GP would **take time to listen**
- **IF** the GP would regard my joint pain **as part of the normal ageing process that one just has to accept**

I would be more likely to consult the doctor about the joint problem if I was in…

Situation A ☐

Situation B ☐

Choice Set 5

Imagine that your joint problem and its impact on your everyday life is changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- **IF** I was experiencing **short episodes of an unpredictable pain**
- **IF** the pain was disrupting **most** of my everyday life
- **IF** I was experiencing **no other health** problems at the moment
- **IF** the GP would offer me **painkillers only**

Imaginary Situation B

- **IF** I was experiencing a **dull aching pain, which was there most of the time**
- **IF** the pain was **not** disrupting my everyday life
- **IF** I was experiencing **no other health** problems at the moment
- **IF** the GP would offer **me a promising new treatment**

I would be more likely to consult the doctor about the joint problem if I was in…

Situation A ☐

Situation B ☐

Choice Set 6 (CHANGE OF FORMAT)

Imagine that your joint problem and its impact on your everyday life is changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- **Short episodes of an unpredictable pain**
- Pain disrupting **most** of my everyday life
- **no other health** problems
- GP offers **painkillers only**

Imaginary Situation B

- A **dull aching pain, which was there most of the time**
-
- Pain **not** disrupting my everyday life
- **no other health** problems
- GP offers a **promising new treatment**

I would be more likely to consult the doctor about the joint problem if I was in…

Situation A ☐

Situation B ☐

We are aware that the choice tasks may become monotonous after a while and feel repetitive. We would like to thank you for all of your effort and energy so far.

We suggest at this point you take a break and come back to the choice tasks afterwards.

Part 3– Your Feedback

Now, please answer the two following questions:

1.Please chose a box, which best describes the way in which you made choices in the choice sets?
(Please put a cross in one box only)

I focused on every feature before making a decision ☐

I focused on most of the features before making a decision ☐

I focused on about half of the features before making a decision ☐

I focused on at most 3 or 4 of the features before making a decision ☐

I focused on just 1 or 2 features to make a decision ☐

None of the above

2. How hard is it for you to imagine these situations overall (e.g. that your joint pain, and its impact on your everyday life is changeable)?
(Please put a cross in one box only)

Not at all hard…….……..…….. ☐

A little hard ……..…… ☐

Quite hard ……..……..……..…….. ☐

Very hard ……..…….. ☐

Extremely hard…..…….. ☐

3.We would like to invite any other comments you may have on (a) the choice questionnaire, for example, the difficulty of the tasks and the layout etc?

……
……..…
……
……
……..
……
……

(b) any other comments on deciding whether to see the GP about your joint pain, for example, are there any other vital factors that we have missed out of this questionnaire?

……
……
……
……
……
……
……
……
…….

Appendix 5b: Pre-planned probes used in preliminary test 3

Specific objectives	Action	Required materials
• Introduce changes made since last session • Refresh choice questionnaire task and 'think aloud' • Outline objectives for today	• Present slides (to summarise changes made) • Refresh choice questionnaire process and 'think aloud' process. The process is the same as last time, with 3 mins at end for group input. • Inform users that we are only interested in some parts of the questionnaire show 'Cover letter and information sheet' on slide and ask them to note if there are any outstanding problems) • Outline 3 objectives for the meeting: 1. Joe 'demonstration of how to fill in questionnaire' section included in questionnaire sent out (compared to a less prescriptive version) see page 357 2. ONLY choice formats 1-5 (as included in the questionnaire sent out to members) will be covered today (and compared to a simplified format) 3. 'Other health problems' factor (minor and major categories) 4. Feedback section: We would like to know if this section is worth including.	Laptop and projector Power point slides
1.Joe 'demonstration of how to fill in questionnaire' section included in questionnaire sent out (compared to a less prescriptive version) see page 357	○ Think aloud and 3 min input at end. ○ At the end ask ' What did you think of that? How easy or difficult was that to do?' ○ After Joe 'demonstration of how to fill in questionnaire' present less prescriptive version (see page 357) And ask ' which one did you find easier to do? ○ Which one would you vote for?	pens double sided questionnaires *1 Dictaphone

Appendix 5b cont: pre-planned probes used in preliminary test 3

Specific objectives	Action	Required materials
2. ONLY choice formats 1-5 (as included in the questionnaire sent out to members) will be covered today (and compared to a simplified format) General objective is to uncover issues of 'comprehension, recall, judgment and response'	○ Think aloud and 3 min input at end. ○ At the end ask 'What did you think of that? How easy or difficult was that to do?' ○ Probe around the levels of each attribute, i.e. were they understandable? ○ How easy or difficult was it to trade against the pain type/severity. ○ Present simplified format) in questionnaire. Ask 'which one did you find easier to do? Which one would you vote in?	
3.'Other health problems' factor (minor and major categories) General objective is to uncover issues of 'comprehension, recall, judgment and response'	○ 'Other health problems' factor (Major/minor exercise) How easy or hard was it to categorise into major/minor ○ Are the choice sets repetitive? ○ Overall length? (more than 16 choice sets would result in respondents cognitive burden?) ○ Should there be an opt-out choice ○ Feedback section: We would like to know if this section is worth including? 1 of this section is based on a previous Sawtooth study into decision-making heuristics used in ACBC	*9 pens *9 Questionnaires dictaphone

Appendix 5c: Information sheet sent to respondents prior to preliminary test 3

KEELE UNIVERSITY

arc Committed to curing arthritis

NHS National Institute for Health Research

PRIMARY CARE SCIENCES

ARTHRITIS RESEARCH CAMPAIGN NATIONAL PRIMARY CARE CENTRE

Address

Date

Dear respondent

Choice questionnaire on how people make choices when deciding to consult the GP for joint pain

We are writing to you as a member of the Research Users' Group at the arc National Primary Care Research Centre. We wondered if you wanted to help us in our final stage of developing our new choice-based experiment.

If you remember, the new study aims to find out about the choices that people make when deciding whether or not to raise their joint pain problem with the general practitioner. We are looking for one volunteer to compare a series of different situations and choose which one they think would be more likely to lead you to consult the GP for your joint pain. One volunteer will be asked to 'think aloud' during this process, and the rest of the group will be invited to join in a group discussion at the end of each page.

We want to ask you about three things:

1) To tell us how effective sections A, B and C (part 1) are in explaining the task and preparing participants for making choices. Little time will be spent on this in the session so if you would look over this section and we will ask for some feedback at the start of the session.
2) To tell us how effective Section C is overall, i.e. the choice questionnaire. We will be focussing mainly on how the factor 'other health problems' work in this design. We have categorised other health problems into 'major' and 'minor' and want to discuss if these are appropriately categorised.
3) To tell us how effective the layout, length and other elements of the design are when making choices.

The final section 'about you' is not going to be a focus for the meeting, as this a well used questionnaire within the centre.

If you agree to help and give us feedback then it will involve you coming to the Centre here at Keele and partaking in a group discussion after a 'think aloud' has been done by one volunteer. The session will last for about 60 minutes and we will ask you about how it went. We will ask you how easy or hard it was to fill in, and how easy or hard it was to make choices about different situations. As this is a new type of survey your advice will be

extremely valuable and we will use your feedback to improve the methods we use in conducting this postal study.

We would be very grateful if you would let us know if this is something you would be interested in helping us with. If you would like to know more about this study, please contact Carol Rhodes at Keele University on 01782 734377

Thank you very much for your help with this research project.

Yours sincerely,

Researcher

T: +44(0)1782 733905 F: +44(0)1782 733911
E: primary_care_sciences@cphc.keele.ac.uk W: www.keele.ac.uk/research/pchs/pcmrc/

Keele University, Staffordshire, ST5 5BG, United Kingdom
T: +44(0)1782 732000 or 621111 W: www.keele.ac.uk

Appendix 5d: CBC partial profile transcript from preliminary test 3

Length 1 hour 3 mins and 39 secs

Interviewer 1 = DC
Interviewer 2 = CJ
Interviewer 3 = CR
Interviewer 4 = MF

Respondent male 1= Research User Group member
Respondent male 2= Research User Group member
Respondent male 3= Research User Group member
Respondent male 4= Research User Group member
Respondent male 5= Research User Group member
Respondent male 6= Research User Group member

Respondent female 1= Research User Group member
Respondent female 2 = Research User Group member
Respondent female 3 = Research User Group member
Respondent female 4 = Research User Group member
Respondent female 5 = Research User Group member

Interviewer 1: Right you're all live now, ok so I'm just going to plonk the recorder in the middle here somewhere, and we have had a volunteer, so, great.

Respondent female 1: Well when I first looked at this I thought oh my goodness what a lot of frames, erm my second thought was I'm a slow reader, erm although my subject was language, erm and then I thought how can I do this quickly, which I thought a member of the public would look at if they were faced with this questionnaire, so, I tried to look at my thought processes as I went along. I notice that the format, the heavy type of certain parts of the sentences and the grey areas were quite a good idea but I found it quite a long time for me to understand the instructions. I think this was more my inability to organise my thoughts from the actual instructions themselves, so, the second page where it takes you through to step one, step two, with the grey areas, I found did put the whole thing in perspective for me, then I knew more or less what I was doing, so, I'm looking now at choice set 1.

Interviewer 1: Sorry XXXX to interrupt you, I just realised, the idea is to go through the whole thing, so to start from the beginning

Respondent female 1: OK so now I'm about to go through the whole thing, that's why I've said choice set 1.

Interviewer 1: Ok, no, Sorry but I mean to start at the beginning, on section B, which is just before the choice bits

Respondent female: OK, so where are we?

Interviewer 1: Page three.

Interviewer 2: Yes, section B, just there, so, if you look at the questionnaire you've got the front page with the picture on and then you've got an instruction page and introduction page and then section, so it's page three, it's section B, so that's the first bit Nica wants to have a look at.

Respondent female 1: Ok, I think I simply read this through erm but what I did before I started the whole thing was to write down what I thought triggered my response to going to a GP and I've actually put those in notes for you at the end.

Interviewer 1: Great ok, thank you. That's all really useful.

Respondent female 1: So these are simply read through and I thought they were very clear.

Interviewer 1: Do you want to do it now, do it for us?

Respondent female 1: What you mean what I've put?

Interviewer 3: If you can read through it as though you have just received this in the post now, as though you've just opened this up, that is what she wants you to do.

Respondent female 1: Well that is quite difficult because it's now a close thing to me.

Interviewer 1: OK if you want to just go through it and read it out loud as if you're doing it now and then the comments you have written you can read out.

Respondent female 1: well A – H I hadn't got any quibble with at all but some of them were ideas I had put down for you at the back and others were ideas that I am in agreement with, I thought you have selected those very well and I thought those were about the right number. Ok, turning over onto this, 'consider the factors from A-H'

Interviewer 1: Ok so there are five factors that are important to me and one of them is on this page, I have selected six factors for this experiment that I am thinking of including and I just wanted to find out if one of them in particular was appropriate and if not, why not sort of thing, so what would be really helpful would be if you could read that one out as if you were reading it for the first time and tell us what you are thinking, erm, is that OK?

Respondent female 1: Yes, I found that again, in retrospect, very suitable for me because that was a strong determinant as to whether I would go to the GP or a particular GP because I have undergone a lot of personal experience of complimentary therapies and had a lot of painkillers and the new treatment idea would appeal so I would definitely include that in your final factors.

Interviewer 2: How about everybody else? Would that point, do you agree with XXXX or? So number D on page three,

Respondent male 1: Yes I do actually because other problems can be contributory.

Interviewer 2: Yes I think it was D, so 'other health problems that may be affecting you at the time', so Nica's put this into the questionnaire, is that something that will influence your decision to consult a GP, do you think that is an important thing that goes through your mind when you're working out whether or not to go and talk to the doctor?

Respondent male 1: yes, it is.

Respondent female 2: But the confusing thing is that it does say at the top, for your joint pain.

Respondent female 1: And also when you get to the GP, they are very single minded and they will focus on one thing as they've got a 10 minute interview and you'll wiz out and you'll think oh but they haven't bothered with the other things I've mentioned, I guess I'll have to make another appointment.

Interviewer 1: Yeah, and is the minor, short lasting minor complaints and terms 'major', is that clear enough?, would you have any problem thinking well where does my problem fit into that group? So is that clear? So, minor complaints would be nothing more than a headache and a more major one would be…

Respondent female 1: Yep, it is…

Respondent male 2: I can understand factor D but I can't see in my own experience, how, say, having diabetes would cause me to go or not to go to the GP about my joint pain and your questions are specifically about consulting a GP about joint pain.

Interviewer 1: OK so you are saying that you can't see how it would influence you.

Respondent male 2: It wouldn't influence me

Interviewer 4: But that's OK cause it might influence other people.

Respondent male 1: I think for most people it will be a psychological factor.

Respondent female 1: Because people are human and we tend to sort of waffle off the main subject and think ooh while I'm here I'll just mention so I think it has relevance, it was in my thinking as I went through the whole questionnaire.

Respondent female 3: But the GP might say well you've come about joint pain today if you've got any other problems then make another appointment to see me.

Respondent female 3: That's the GP's choice though isn't it, really?

Interviewer 1: Yeah, ok, yeah, great, so that's five minutes is it? OK my timekeeper is telling me to hurry up, so OK do we want to move onto the next section, is that ok Hilary?

Respondent female 1: Are you happy with the other points then?

Interviewer 1: Yeah I think so, I mean as long as you are all quite happy that A-H are all quite relevant then yeah.

Respondent female 1: Right, again so I filled this in with circles that are personal to me which I'm sure you don't need at the moment but I have put that I thought the weakness of multiple choice, I tend to compare re number choice especially the last three FGH which are similar and I put a bracket around E and F because they can be fused together, but it's always a weakness of multiple choice. You tend to compare one with another and the choices themselves are limiting rather than a person themselves writing a paragraph but I realise not everybody can write a paragraph.

Interviewer 1:OK, right

Respondent male 1: Right, are we on page 4?

Interviewer 1: It's interesting you say that about E and F because originally they were linked together, now we've changed them.

Respondent female 1: They could be fused together and that would reduce you to G wouldn't it instead of H?

Respondent male 1: Stuart phoned me up this morning and said that he wouldn't be able to attend and I think we both picked up on the same thing, that the actual question, not the ABC etc. but at the top could be confusing
Interviewer 1: Ok, in what way?

Respondent male 1: well I know at the time I wasn't feeling very well but you've got to circle number 1 to least important to 10 most important, I thought how do I put a circle around. It should be stated one circle for each question.

Respondent female 1: I'd put, instead of putting a number, you should put please circle one number.

Interviewer 1: OK great, thank you that's really really useful.

Respondent female 1: Sorry to go back to this but looking at G and H, the GP's manner is G and the attitude to joint problems is H, I think those two are likely to be fused together on second thoughts, having gone through it a second time.

Interviewer 1: To some degree that is true isn't it.

Interviewer 2: It might be useful for Nica to say what she means by the general manner, than what she means by attitude, if you, is there a distinction?

Interviewer 1: Yea I mean the manner is more about his/her consultation style in general, whether they listen or whether they don't and the attitude is more specific to joint pain so it's the attitude to the actual health problem, joint pain and it's kind of a subtle, cause attitudes are hard to detect really, erm, so I suppose the manner is what is happening physically and the attitude is more hidden, but there is a distinction there but I can see how there are cross over's.

Respondent female 1: It could be fused to make the whole thing shorter and easier to fill.

Interviewer 1: SO, if you are going to fuse them together what would you put? Which one would you go with or would you just choose one?
Respondent female 1: erm, I think I would fuse together E and F treatments and the further investigations could be combined quite nicely, it depends how you phrase it in both cases and I would fused together G and H, using the GP's approachability and the GP's attitude to joint problems. Approachability I know is an old fashioned word but it's one I think older people will relate to and most of your fillers in will be mostly the upper age group I imagine.

Interviewer 1: Ok, that's good, great.

Respondent male 1: The other one on the top of page 4 is I've also noted that on page 28 of the original document (where it says part 2) it is much better worded than the top of page 4, so you are better off using page 28-24.

Interviewer 1: Ok great, so we are going to swiftly move on now.

Respondent male 2: Before we move on can I say that E, F G and H are on expectations aren't they? Because none of those will have arisen until you've got yourself into the consultation and what you're asking is whether somebody is going to be bothered to go for a consultation so the sorts of treatments the GP can offer you is really what you think the GP is likely to offer, that is what you expect, and what you expect the manner to be and what you expect the attitude to be rather than what it is?

Interviewer 1: That is why this method in itself is confusing because it is hypothetical so in a way your expectations will come through, but in a way it's that idea of suspending your disbelief, and suspend your own experience for a bit and pretend that actually these things are available.

Respondent male 2: But you can't, I was stuck, when I went through it carefully, with the fact that I would not know what the treatments are on offer until I have been to the consultation, and what you are asking is whether I would go for the consultation, and similarly for F and G and H.

Interviewer 1: OK, is it worth? I mean how would you make it simpler for yourself? i.e. would you say something like, because often in questionnaires the more you say the more confusing it can be so.

Respondent male 2: well instead is likely to 'expect to offer' or 'think might be offered' something like that.

Respondent male 1: Put it as more of an expectation than as a..

Respondent male 2: And I'm not sure of the value, at this stage, of separating E and F and G and H.

Interviewer 1: OK, right, thank you

Respondent male 2: Cause to my mind when thinking about going to the GP I would merge together manner and attitude I wouldn't think about separating them, nor would I think about whether it would make a difference to me that I thought treatments or referral might be offered, something might be offered, so I would tend to think of those as being together.

Interviewer 1: Ok that might help actually to reduce it because we really want to try and simplify as much as we can, that's definitely something we can think about and revisit.

Interviewer 2: What do other people think about that? Anybody else?

Respondent female 2: Well you go to your GP expecting him to tell you what is available, cause you don't know so that's not a reason why you would chose to go to the GP is it? You are hoping, you are in pain so you want something that will ease the pain?

Interviewer 1: Ok so is it the wording because I suppose what we are trying to say is he will offer you this, he is going to, so it's a reality, I mean its hypothetical but it is a reality.

Respondent female 2: Or that he 'will be able to offer you' some treatment.

Interviewer 1: OK

Respondent female 2: Because you don't know what the GP will be able to offer you, you know physio, painkillers, whether the painkillers need further investigation or what,

Interviewer 1: Ok, great, we are going to have to move on, so, the next page, page 5, we are going to skip as well, I mean we are going to skip bits but if you have any major outstanding things then do raise your hand and say, but otherwise we are going to focus on these main areas. So, the next bit is page 7, sorry I mean page 9.

Respondent male 1: Sorry can we just go back to page 5, well my first impressions were that it was thick and daunting and in the last paragraph on page 5 you have got, it suggests you take a break period, now that, I mean its thick in any case and if I'm going to need a break after it, that's just proved it to me that you're expecting a bit much.
Interviewer 1: OK,

Respondent female 2: If she shortens the whole questionnaire she won't need that bit so that's off anyway.

Interviewer 1: OK, great, so page 9 then.
Respondent female 1: What about page 6 cause I did say I had a bit if difficulty deciphering the instructions, it was these grey loops, is this an indication of Joe's thinking, these grey loops. I couldn't work out what they are.

Interviewer 2: OK, so the section we are on is 'how to complete the choice questionnaire' is what we should be looking at. SO If you're looking at the new one that Nica's given out today it's on page 7.

Respondent male 1: I'm just looking at Hilary to see what page you're looking at (laughs).

Respondent female 1: I'm looking at the original, I find it very confusing that I've been issued with two lots of information, because I have already annotated the first lot of information, so could I request it's the same numbers you put on next time so that we know what we're doing, cause I keep going back to my own one.

Interviewer 1: yeah, ok, sorry about that, yeah, ok.

Respondent female 1: Erm, I couldn't work out what these grey loops were, was it an indication as to the choices that Joe was making?

Interviewer 1: Yeah they're his priorities really, yeah.

Respondent female 1: Because the instructions are actually, or the way he thinks actually follows that step 1, step 2 and step three, so I felt really it was in reverse order, or it should be alongside it, also this grey area is explained at the bottom, it would be helpful to me if that instruction 'please remember that the greyed out boxes' was at the top.

Respondent male 1: yes, that should be on every page I think as well

Respondent female 1: It was also confusing that the second lot of greyed out ends at 'ageing process that', so I'm looking at that as an ex English teacher thinking…

Respondent female 1: there should be something else there.

Interviewer 1: Yeah that's cut off.

Respondent female 1: It's missing. If it's meant to be parallel or identical it needs to be kept as identical, I mean the point of English, its stuff that people would pick up on I think.

Respondent male 2: that bottom statement is not correct is it? Because the lozenge one is shadowed on the one side and not on the other and yet you are saying the grey areas mean it is the same across both situations.

Respondent female 1: The ellipses are confusing, if the ellipses were left out it might help, it's confusing to use that and those greyed out at the bottom.

Interviewer 1: Well, that's why what we've done, actually we'd like to include a demonstration but I'm not sure how to do it, because originally as you said Hillary, the steps were all on the same page but it was really confusing so I changed it and put them on a different page and so now that is still not quite right, so I suppose really the question is, er, is should we include it? Is the question, is it too confusing? Is it easier to just give a brief summary, erm the new piece I gave you today, if you look to page 9 in the questionnaire I gave you today, I have included an alternative, so, it's just a short paragraph about how people fill in questionnaires and then leave you to do it. Today what I'd like to know is which approach would you prefer? Would you prefer a more in depth, prescriptive approach or a less prescriptive one?
Interviewer 3: So, we'd only have the paragraph on page 9? Do we prefer that or do we prefer page 7 and 8.
Respondent female 3: I found page 7 and 8 very confusing.

Respondent male 1: yeah.

Respondent female 3: because of the steps, I'm thinking are you supposed to go down, are you supposed to go step 1 and then you go and I just thought it was just so confusing.

Interviewer 4: Different people do it different ways, some people look at a group, others will, and that is why we have said that different people do it in different ways so just go ahead and do it.

Respondent female 3: I don't think you need, personally, I don't think you need that bit.

Respondent female 1: It's either that or that and that to me requires you to be an academic and that means you are directing thought.

Respondent male 1: The paragraph is better yes, it's shorter, easier.

Respondent female 1: It's a better idea as the gentleman over there suggests, leaving it free for people to use their own minds and which way they read that.

Interviewer 2: What do you think?

Respondent female 3: Yes I can see the points that are being made, but it's not easy to find out (demonstrate), in reality its totally, when we read something that's written like this is and it's just going down in steps it does get a bit confusing.

Interviewer 1: So perhaps the approach of giving just a little information and then throwing you in a bit really, would you say that is the better approach?

Respondent male 1: **yes**

Respondent female 1: yes

Respondent female 1: You could eliminate that, because that would be an off put to some people.

Interviewer 1: There are too many ways of answering it to be prescriptive about it, OK; does everyone agree with that then? Does anyone disagree with that?

Respondent female 4: I think you might also get people say 'oh I can't be bothered with that' you know what I mean, who just couldn't get their head around it at all so therefore, it's not helping them.
Respondent male 1: I didn't even read it to be honest with you.
Respondent female 4: See, I rest my case my lord (laughs)

Interviewer 3: I did that I didn't read it to start with and then I had to go back and read it cause I thought it was going wrong.

Respondent male 1: I just looked at it a thought oh I'll have a look at the questions.
Respondent female 4: No brownie points for you next time mate.

Respondent male 1: No, sorry.

Interviewer 1: No I think that's the best decision actually.
Respondent female 1: Are you happy with that Nica?

Interviewer 1: Well, yes I am, there have been a lot of these done, you know discrete choice experiments, and I've seen that no one's really given a demonstration and thought god you can't just leave people to get on with this, but actually it might well be the best approach.

Respondent female 1: Yes, you can be too helpful, so people get off put by it.

Interviewer 2: Shall we just look at the wording of this?

Interviewer 1: Yes, that would be great, so Hilary if you want to just have a read through of page 9, just read it out for us if you would.

Respondent female 1: **You'll be charged for this (laughs)** '**People have different ways of filling out the choice questionnaire. Some people compare the factors between situations and others will look at the situations as a whole and try to compare the situation against the other one. Others simply choose the factors important to them and overlook less important factors. It is entirely up to you how you make your choices'. That sums it up very nicely that last sentence I think.**

Interviewer 2: **What do people think about that?**

Respondent male 1: **Again I think with actually knowing what I'm supposed to do in any case, it's pretty straightforward, I'm worried about people, whether people will understand what has got to be done, I mean I can understand but will others?**

Respondent female 2: I think that last line is important 'it is entirely up to you' how you do it.

Interviewer 4: You could actually put that as the first line? To reinforce it right at the start.

Respondent female 2: You could actually.

Interviewer 1: Ok great.

Respondent male 2: I find the statement itself helpful, what bothers me is that with such a variable way of working out which box to cross, what is going to be the value of the findings?

Interviewer 4: But I think I mean I did it as a hypothetical person who doesn't have problems, and I found that when I had the choices I did it in different ways for the different scenarios. Sometimes I chose one factor, sometimes I looked at the whole group and as a hypothetical person I found that I did have completely different, and I don't think it shows that one is right or wrong.

Respondent male 2: I'm not questioning whether it's right or wrong, I'm questioning in my own mind that if for example on this one here, somebody decides that it is the first thing that matters so they tick the box and then somebody decides it's the second factor and third and somebody decides it's the fourth factor, this is not a heterogeneous response, but simply a choice between A and B, what is going to be the value of the findings when some people are and some people are not taking account of everything that has been raised.

Interviewer 4: That's OK cause what we'll find through the analysis, cause we'll do it in hundreds, you know quite a lot of people, we'll find some people do it in this way, some people do it in that way and that's fine. There isn't a right or wrong answer.

Interviewer 2: So that's an issue about the actual methodology of this type of project in how people make their answers and so it will be different and it will be variable but I think what you are trying to say is that doesn't matter, the analysis that you'll do, for this project, for the way this analysis is done, it doesn't actually matter.
Respondent male 2: OK

Interviewer 4: OK, it' would be like saying to me are you going to vote conservative or labour, some people will vote labour based on foreign policy, some people will vote based on pensions or whatever, people will choose in all sorts of different way, it's only the choice

Respondent female 1: My choice also changed as I went through, my way of thinking and selecting that choice changed as I went through because I thought in this questionnaire, she's playing games with me here, so you couldn't really be. One is I was choosing physio out of that and then I was rejecting physio because something else had cropped up so I thought there is something subtle at work here, so I'm going to have to adapt my thinking and be flexible with my thinking, so I what I've done is I've written the way I think across every situation and I said to you I'm going to tell you that but I can't get started, and after a while you will say end of time.

Interviewer 1: OK, so we're ready

(People talking at the same time)

Respondent male 2: The choice of how people decide how to go or not to go to the GP and I'm wondering if people are using the question in so many different ways whether in fact that is going to illuminate how the people come to make their decisions.

Respondent female 1: But that doesn't matter, that's their decisions, they are the researchers. We are the.

Respondent male 2: Is this instrument capable of doing that?

Interviewer 4: We're trying to see if it's better than the standard way of answering questions. People have found this way is better.

Interviewer 1: I think it's time that Hilary is going to let rip, so we are going to listen to Hilary now.

Respondent female 1: Are you ready to flick through folks, Are your fingers on choice set 1, all copies preferable. Right I thought oh goodness I'm never going to get through this so I thought I'm going to have to compare A with B, so I chose, and just out of interest I chose situation B, that's a personal choice everybody else might have chosen A but that won't bother me.

Interviewer 1: Hillary, would you mind just?

Respondent female 1: Could I just rip through showing them the way I thought first please, so move onto choice 2, I read down using arrows only because my motive here was speed, I was getting a bit of fed up this and that was on frame 1, erm, moving onto choice set 3 and by the way I chose situation A on 2, but that's immaterial that's just my personal choice from what I've experienced of the medical profession in the past and the help or the non-help I will get. Choice 3, now I am marking the outstanding factor which I have marked out for you, erm, I chose situation B. Choice set 4 there were four points with no grey blocks out, so I read vertically, so I was influenced by this grey blocking which I ultimately decided was a bit of a waste of time. There were four points so I read vertically and I chose situation A, erm and at the bottom of A I wrote 'a bird in the hand' i.e. physio. At the bottom I have written the last point, asterix, if the GP would regard my joint pain as a normal part of the ageing process, out, the second sadly.

Interviewer 2: What do you mean by that Hilary?

Respondent female 1: It cancels out,

Interviewer 2: It cancels out your choice on A?

Respondent female 1: It cancels out the fact that the GP would regard the joint pain as a natural part of the ageing process, because in my mind if he thinks that he is not going to do anything. Moving onto set 5 erm, we've got to keep it moving, I read down and across because I decided the questionnaire was getting more complex. The gentleman in the corner is nodding, I'm sorry I don't know your name.

Interviewer 4: Martin

Respondent female 1: OK thanks Martin. Erm I chose section B, and under A I've written that most people can buy their own painkillers, the pivot point there is that the 'pain is disrupting most of my everyday life'. Most people can buy painkillers I've written, and on B, I've written the grass is greener and I went for B, for some reason, 'promising new treatment' that was the carrot dangled before my nose. On Set 6 I went for situation A, erm and I queried here 'not much difference in the understanding of most and some' if the pain was disrupting 'most' of my everyday life and then opposite you've got 'some' of my everyday life, in my way of thinking, there is not much difference between those two words, most and some and I've chosen A.

(people talking over each other)

But I must say that my pain has been going on for so long and is understood that I tended to disregard the pain throughout the whole of the filling in of this questionnaire. I regarded the pain as understood whereas I feel to yourself you are making it a prime issue, fair enough, that's what you want to look at.

Interviewer 4: yes exactly that is a focus area for me.

Respondent female 2: On number 7 erm the decisive factor for me is if the GP would offer me a promising new treatment, have I got to stop?

Interviewer 2: No I think because you have gone through it you could carry on, then this would really help, since you've spent the time going through it. Some of the issues about how you are making decisions are exactly what Nica and Martin want to understand whether you are doing this task as they want you to do it, as they want people who are going to fill this in to do it, so it's really good stuff so please carry on.

Respondent female 1: I was trying to look at the way I was thinking as I was going through it I wasn't sure if I had succeeded or not but the decisive factor for me was, and obviously when you get to this point you have got to read and compare, because you've got your two greys at the bottom which to me was material understood, so the decisive factor was if the GP was promising me a new treatment. Again it's the grass is greener on the other side attitude, and set 8 I chose situation A. this demanded a bit of thought because you've got two grey fill in here, erm, which are identical are they not?

Respondent male 3: Yeah you've only got two choices then haven't you

Respondent female 1: She's forcing you to look at pain, which is good as she wants you to test pain, 'if the pain was severe, moderate' erm, and the other one 'if the GP would offer me medicines' erm and 'if they would encourage lifestyle advice' well lifestyle advice can be got from any pamphlet in any hospital. Of course being a complimentary medicine woman, I went for the complimentary therapist. Then I thought it was nice that we'd take a break because I thought people would be collapsed with that, I nearly collapsed with over thinking at that point, I had to go and lie on me back for 25 minutes after that and then I'll put the bill in later, so choice set 9 I've chosen A, I felt there are three decisive factors: no. 1, no. 2 and no. 3 and I've ticked them all for you on my sheet, set 10, the deciding factor was in B number 3, if the GP would offer me a promising new treatment so again I'm continuing this same line of thought and I want something new from the GP, I've been umpteen number of times over the years since I was 30, I'm now 66 and I've got nowhere, so I want something new, so it depends on your previous experiences how you fill it in really. Choice set 11, the deciding factor was the second one on B, 'if the GP would refer me to a hospital specialist/rheumatologist/orthopaedic doctor or pain clinic' I've tried the pain clinic, had a TENS machine, I had to take it back, and all they did was talk to me at the pain clinic. If they could send me to a rheumatologist or an orthopaedic doctor, which I saw in the 1970's I would be thanking Allah, erm, this one if the GP, last one on B, 'if the GP would regard my joint pain as a natural part of ageing process that one just has to accept' erm, I put here, that obviously needs treatment because its part 2, I don't quite know what I mean by that.

Interviewer 1: Part 2, do you mean the wording of it? As part 2 rather than of, is that what you mean?

Respondent female 1: I'm comparing now with A (reading over) OK, I mean that I obviously need treatment even though the GP has got this attitude, that's my little anecdote added on.

Interviewer 1: Right OK.

Respondent female 1: So I chose B, on 11, and obviously you've got to look at these in a comparative way, horizontally and you've got to look at these vertically now, so I thought it was quite clever, how the questionnaire developed at this point.

Respondent female 1: On 12, erm, I've gone for A, we're looking at three things and I've put spectacles, meaning this is important by the third point in A, meaning that the GP would offer me a promising new treatment, so all the time I was flashing about choosing what was new being as I've had the problems such a longstanding time, but I thought it was cleverly devised. The first half could be very much compressed because it was so repetitive. My comments at the end were, er, how set 14 introduces a completely new treatment, so I was surprised that the thing had gone on with much sameness until 14 when you offer an injection, so it takes a complete new turn there.

The paper version is more appealing to elderly people I think. I, my personal view, found there were too many frames; the frames were too repetitive for my concentration (which is very poor). I would suggest 10 frames. Frame 14 introduces a new temptation, which is an intra-articular injection and I didn't know what that was, you must explain that to a lay person, the print size and layout was very appealing, I could not always see the sense of grey blot outs, especially when they were the same. That's my summary of it.

Respondent male 1: That was the idea of the blot outs.

All Respondents: (clapping)

Respondent female 1: Sorry for starting at the wrong point and getting confused easily.

Interviewer 1: Not at all, I'm sorry for confusing you with all the different papers.

Respondent female 1: well you can take my number and you can ring me if you need anything else, I don't mind at all.

Interviewer 1: Ahh thank you Hillary.

Interviewer 2: There have been a huge amount of issues there that Hilary has picked up on, so, I think maybe if we go around the room, because everyone else has had a chance to think about it while you have been going through it Hilary, so, can we just pick up on what, maybe 1 or 2 points were as we went through. Can we maybe start with XXXX, is there anything in addition to what Hilary has bought up or would you like to reinforce or provide an opposite view to some of the things Hilary has highlighted as she has gone through the questionnaire.

Respondent male 2: I feel I've raised already most of my concerns but there is another one which is in my case, with 37/8 years' worth of having to deal with this chronic pain I'm not in a position to judge whether my pain is interfering with my daily life because you perpetually adapt minute to minute, what I am better able to judge and what might be more clear cut is how much I am concerned by the particular pain I get, if it is extremely familiar, whether it is high or low, you know how to ride it. If it brand new or relatively new, whether it is high or moderate, you are more likely to pay attention to it, to assess whether that warrants going to a GP and that's not the kind of thing you are asking here, but it certainly would drive me way beyond this stuff about how much is it affecting my life, it is how concerned am I, whether it is intermittent, regular, high or low and that to a large degree is how familiar is it to me in my circumstances.

Interviewer 2: So, is that, would you say that is not covered by the type of pain we've used 'dull aching pain' and 'unpredictable' pain.

Respondent male 2: I couldn't see how it was covered at all and what I was telling you to do in the long version that you sent out, I rephrased for my benefit, this disrupting everyday life, into how much would I be concerned about whatever was the choice above, and I found and felt better able to operate your choices that way.

Interviewer 2: OK

Interviewer 4: Just to mention one of the things that we were trying to do is to keep the description very short, which is why in one of the examples that I think Nica's going to talk about in a minute that is to deliberately cut down the words to make it as simple as possible, and clearly the issue you have raised, is you know that we are trying to play off these two issues, complex factors but keeping the description very very short.

Respondent male 2: I think you could replace that with something briefer about the level of concern rather than the level of disruption, in part because I am in no position now to judge the level of disruption, if it starts to interfere I just swap and do something else so is that interfering with my life or not, my answers unknown to me, my wife might be able to tell you but I cannot and I'm the who would have to answer the questionnaire.

Interviewer 1: OK would you say something like 'I'm very concerned, a little and I'm not concerned'

Respondent male 2: erm, how concerned was I about this pain level or this pain. I've just got another thought, who exactly is being given these sets of questionnaires, is it people with chronic pain, is it people with joint pain or is it the public at large?

Interviewer 1: its people that have current pain either hip, knee or hand, the severity is variable.

Respondent male 2: Is it chronic or not?

Interviewer 2: It could be, we kind of use a working definition that research uses of chronic pain is more than three months so there will be people with chronic pain, there will be people with non-chronic pain who've had it a short period of time, people that have had severe pain, less severe pain, so it will be a mix of people with mixed characteristics of their pain, and they might have pain in one joint, two joints, most people in this age group will have pain in more than one joint, so it will be a mixed population who reply.

Respondent male 2: My phrase would be 'how concerned was I' or 'I am concerned and I am not concerned'

Interview 1: Ok great, thank you for that, just on that note can we look at page 15 of the new copy, because we are talking about simplifying it and on page 15 we have a simplified version which will really just cut out a couple of words out at the beginning, so instead of saying if the, we've cut that out and just written what the circumstance is. I wanted to know from you guys if that is a better way of doing it or is that.

Respondent male 2: what page are we on?

Interviewer 2: We are on page 15, choice set 6.

Respondent female 1: You want the wording looking at Nica?

Interviewer 4: It is about whether the shorter version is OK, by missing out, just to have a brief description.

Respondent female 1: I think that is much easier, people can't cope with a lot of verbiage, particularly with the onset of technology. I've got to get my stab in against technology haven't I really?

Interviewer 1: Does everyone agree with that? Is that Ok?

Interviewer 4: The concern we had back from last time was that 'we can't imagine that situation' which is why we have put in this new version, 'if I was experiencing' so that's to reiterate that, but if people are happy to accept that this is a hypothetical situation then we can drop quite a lot of the words and make it simpler?

Respondent female 1: I would compress your vocabulary as well to reduce your number of frames.

Interviewer 1: Great OK,

Respondent female 1: It will help with the number of takers or the number of people who can stick it to the end of the questionnaire.

Interviewer 1: Exactly. And we have got imaginary situation A and imaginary situation B, so if that's sufficient we can get rid of the 'if'

Respondent female 1: yes

Mixture of Respondents: yes

Respondent female 1: Are you going to put the IF in at all in the first couple of frames or not?

Interviewer 1: I don't know, maybe that is the solution maybe?

Respondent female 1: I think yes I'd do it on the first frame and then put a note at the bottom, 'IF applies to the other frames'

Interviewer 2: What do people think?

Respondent female 1: If you want to protect yourself against criticism but that is..

Interviewer 4: So your bottom line is the shorter the better.

Respondent female 1: Yes. Some little pictures of people in pain at the side (laughs) sorry that would be biasing people towards think this is dealing with pain, I take that back.

Interviewer 1: SO if we put IF… here would that be OK, so there is an IF…

Respondent female 1: or your experience then semi colon, two dots, but then people don't always know what a semi colon is these days. It's difficult with language changing so rapidly these days.

Interviewer 1: OK, excellent, yea, thanks.

Respondent female 1: I think page 15 looks very good really but again if I was looking at that for the first time I'd want some instructions like, IF I experience, so you could put that at the tip IF… or in capitals or in a balloon or something to get their attention, 'If you experience the following' but as Martin says leave the way people read it horizontally or vertically.

Respondent male 1: I don't think that's going to make much difference, just taking one word out, I know it's the beginning of each sentence but it's not going to shorten the document.

Respondent female 1: But it's so repetitive it clogs up the whole sentence, but you want to focus on the essential in the sentence, as you move through it because it's such a long lengthy questionnaire.

Interviewer 1: yes OK.

Interviewer 2: So we are now looking on page 14, choice set 5, so you've either got IF, IF, IF or at the beginning of each sentence or on 6 you've just got the sentence, so that's the, so which one did you think Adele?

Respondent female 3: I think this is better, when you do compare there is an awful lot of reading there and if you've got a lot of questions to read as well, you know like, probably like Hilary says, the first question its fine you can do that in full and then I don't think after that you'd be fine.

Respondent female 1: And repetition as an ex English teacher drives me mad. I think well why don't they start the sentence in a different way, so better to me to lead the IF off and leave a brief instruction at the top, if you need it that is, and maybe you will need it otherwise because people will think I've got two blanks here, what do I do with it? Somebody will think, who's not use to doing questionnaires.

Interviewer 2: No it's fine. So the intra articular injection that comes in at about choice 8 or something that should be explained at the beginning, in section B on page at the beginning is supposed to describe the levels, so that should probably go in to there shouldn't it?

Respondent female 1: While we are asking, is someone going to be present with the person or are they going to be on their tod?

Interviewer 1: This is going to be postal.

Respondent female 1: The instructions are going to have to be very clear then.

Interviewer 1: Yeah, indeed.

Interviewer 2: So on that point Hillary do people think when you read the first instruction pages we've talked about the 'JOE' bit and that may need to be changed, A is the introduction, and section A so looking at pages 1 and 2 comparing one and two, how do those feel to you in setting up the questionnaire and what you need to do?

Respondent female 1: we are looking at one and two now?

Interviewer 2: Yes, section A so where it says the introduction to the questionnaire and where it says instructions for filling in this questionnaire.

Respondent female 3: I'm wondering would it be worth putting down another section at the end to reveal after say 3 months or 6 months, so it just ends there, there is no come back, would it be worth sending another questionnaire at 3 months to see how you've got on with the physio, whatever you had. So you think, oh there is somebody there looking after you and it's to some avail, whereas you can just fill this in and you know..

Respondent male 1: yeah.

Interviewer 1: nothing's going to be done with it…

Respondent female 3: well exactly, that's it, when you think there is going to be a follow-up they might think more seriously, put more thought into how they filled the questionnaire in, because there is going to be a follow-up and they want to know how you are going to be in say 3 or 6 months' time, whether the treatment has been of any help.

Interviewer 1: I think that's a question about erm why researcher wants to do this and what the research question that researcher has, cause if she wanted to follow-up people to find out if their pain had changed for example, then you would need to do another questionnaire later on but that's not exactly what she wants to do, she just wants to find out what makes people go to the doctor and so you can do that just by one questionnaire because she doesn't need to find out if pain changes or if their healthcare changes, she just wants to know what makes people go or not go, so in a way…

Interviewer 1: Possibly, there could be a sentence saying the results will be useful for.

Interviewer 2: We have a few sentences like that which just say, how it will help us.

Interviewer 1: To say something like the aim is to improve healthcare. I know what you mean.

Respondent male 1: I think with it being such a large thing, from past experience I and no doubt other people felt the same, that you've filled in questionnaires and sent them off and that is it. Forgotten. Now, if there was some kind of feedback to them.

Interviewer 4: We could feedback the results couldn't we?

Respondent male 1: What I think it is people want some kind of feedback, even a shortened version of the results.

Interviewer 4: Yes we can do that.

Interviewer 1: Because in the centre traditionally because you know we send out to 20,000 people and it's difficult to do another mailing for 20,000 people of all the results, so but this is a smaller

survey so we could build that in, if we can put it a request in for funding into the project grant, that's something we can put to the team.

Respondent male 1: People will probably be more forthcoming with filling it in, knowing that…

Interviewer 1: Will we put that at the beginning or the end?

Interviewer 2: At the beginning, yes.

Respondent female 3: I'd put on that section A, just looking at it for the very first time, erm when it tells you about the 16 choice tasks, I've just put quite confusing, obviously once you start doing them you understand but just reading that for the first time.

Respondent female 1: I crossed that out, to me, I would stop at 'consult the GP for your joint pain' that you've got in heavy print. Now the factors are really reasons in your thinking the way you've devised this, I don't think the general public is concerned about that, I think they are concerned with getting the questionnaire finished, like I sped up. So, I've crossed the factors to the rest of that paragraph, crossed out the second and third paragraph and then 'it is important to remember that these situations are hypothetical', I don't want to sound condescending but some of the public will not know what hypothetical means.

Interviewer 1: Ok, so 'imaginary' is better.

Respondent female 1: That's right, we ask you to put yourself in their situation, fine, (reading under breath) 'there are no right or wrong answers', that's important and you could actually highlight that 'we are interested in your own personal preferences based on how you might feel in each situation. Please take time to consider'. To me there is too much here, I would go straight into the questionnaire.

Interviewer 3: So include less information..

Respondent male 1: I've actually got down for paragraph four.

Respondent female 1: I was tired by the time I got down to paragraph four; I thought are they going to pay me for this? Well that's what the general public will think, they don't do something for nothing these days, and this is also about life as it is today.

Respondent female 1: I used to read essays like this and I'd get exhausted to write them up you know what I mean, take out your main threads that you want.

Interviewer 1: Ok, thank you, that has been really helpful, thank you all. Erm the only other thing is the information sheet and cover letter, cause they are going to be included as well, but they are actually quite straightforward, what I'll do is I'll give you, is it OK to give you a copy for you to have a look at and then you can see what you think of them, but I'm not going to overload you with it now because there has been a lot of information today.

Interviewer 3: What about the feedback section?

Interviewer 1: Hmm I'm not sure..

Respondent female 3: Can I ask you, are you going to put this other stuff at the back? Section D in the same questionnaire, whoof! see I'd treat that as a separate document, and I thought some of the points which hinged on what XXXX (Respondents male 2) was saying but didn't deal with it completely, as that's the limitation of the questionnaire method, you are making a choice which somebody else has really steered you towards or limited you towards.

Interviewer 1: Actually some of this information won't be necessary, because we are sampling from existing records so we will have a lot of this information, especially the how to cope stuff.

Respondent female 3: I think this is vastly improved though from the first version.

Interviewer 1: That's good news.

Respondent female 3: Section B is a separate document though surely, you've worked so hard that's what you've done.

Interviewer 2: the remaining thing then is part 3, your feedback, do you want to just ask about that then, because Nica's wanting, because this is a new way of doing it, asking in this way, in our field healthcare there are not many examples so Nica was wondering whether to put this bit in to everybody asking about their feedback on doing this type of thing. What do you think about having that in the questionnaire?

Respondent male 3: What would be the value of that page?

Interviewer 1: Well the first part three about finding out how they make decisions may help us in understanding the way they have answered their choices really

Interviewer 2: which is related to your point.

Respondent male 3: I can't tick one of those boxes because in some I did one and some I did another and some I did another again. I have used at least about four of those throughout the system, so if I tick one box, what will be the value of that to you? It could lead you to feel like you've got a secure result when you haven't.

Respondent male 1: Or please choose which boxes so you can tick more than one choice, so you're not nailing it down to one choice.

Interviewer 1: Yeah, so please choose which boxes, yeah OK. The other purpose this serves is, that the next question..

Respondent female 4: Sorry where are we?
Interviewer 2: The one that Nica has given out today, Nica's wondering whether she should ask everyone who is filling gout the questionnaire whether she should have a feedback bit in the one she sends out.

Respondent female 1: Well it's a good idea, if some people will want to and other's wont. I think it's a good idea myself but leave it as optional feedback.

Interviewer 1: the other bits really just give you a chance to express yourselves, if you really hated it, or if you think all of the attributes are just pointless, it's really serving that purpose for people if they fell a bit frustrated.

Respondent male 2: But what would you do with that information cause that's the point of it, to make use in some way, of practical value, so what would answers give you?

Interviewer 1: I suppose it would help to explain the response rate if it was really bad.

Interviewer 4: cause this is a research study and most studies never do anything like this, they just say here are the results, we are trying to go a little bit further and say well we would change the method for these reasons?

Respondent male 2: So your findings might be that this is not a very good way to go about this.

Interviewer 1: exactly, cause that's a finding within itself.

Respondent male 2: Is that publishable?

Respondent female 1: You could ask for two comments couldn't you, at your discretion, about how you felt about it?

Interviewer 4: There is evidence which suggests this approach is better than traditional ways and we want to know if it is like that in health, it may not be and if it's not then we'll say that.

Respondent male 2: I'm not saying that, my question as always is an honest one, the implication is what are you going to find from this?

Interviewer 1: It's a very valid question, yes.

Respondent male 2: Therefore do you want the people to spend the time giving you the info.

Interviewer 4: I think the idea of making it optional is a good idea.

Interviewer 1: Yes, so shall we put optional section.

Respondent female 1: But you're obviously trying to find out these points that I suggested should be obliterated really aren't you, the nearest answer you can give to XXX until you have gone away and slept on it.

Interviewer 1: (laughs) yes absolutely.

Respondent female 1: I think you're a stoic, I think you're really tough, to stand up to the criticisms of this group.

Interviewer 1: No it's really helpful, I don't see it as a negative thing, and I see it as a really really positive thing. There is no point in..

Respondent female 1: I would have felt destroyed being a sensitive soul.

Respondent male 1: I think after the computer one, even Alan said that cause we literally destroyed it.

Interviewer 2: If you think about it though the reason why we are so passionate about PPI is that if you think about Inca's project has changed as a result of your input, to the design of her project, it's a new area, it's not been done much in healthcare, we didn't know how to do it, the team were really like 'we've got to start somewhere' we made one choice, let's go with online, do it, that bombed, you didn't like that at all, go back to the drawing board, let's do a new questionnaire, new format and you can see how your input has really influenced Nica's project, had a real input to the design of her study.

Interviewer 1: It's really helped.

End of tape

Appendix 6: Total list of attributes considered in developmental studies

Attributes		Attributes cont	
1	Pain Severity	20	Disability/Independence
2	Other illnesses/health problems	21	Age/gender/race of GP
3	Curability	22	Affecting work
4	Doctor's access to medical notes and knows you	23	How would I like my GP to be
5	Thoroughness of physical examination	24	Distance of GP from home
6	Doctor's interest in your ideas about what is wrong	25	To permanently judge pain on daily basis
7	Time to get appointment	26	Consistency of opinions.
8	Pain compared to usual	27	Expectation of painkillers and we have to live with it
9	Likely pain relief	28	General conversation with GP to X ray or acupuncture
10	GP's manner	29	Requirements for physiotherapy
11	GP can offer you	30	Understanding and advice from GP
12	Amount of information	31	Imperative for full understanding
13	You always see	32	Positive help
14	Pain compared to usual		
15	Pain type		
16	Impact on everyday activities		
17	GP attitude		
18	GP follow-up services		
19	GP treatments		

Appendix 7: Demographic and socioeconomic characteristics of 1563 eligible participants

Characteristic	Number (%)
Age (years)	
50-64	464 (30)
65-74	631 (40)
75+	468 (30)
Female gender	884 (57)
Marital status	
Married	1073 (70)
Separated	11 (1)
Divorced	101 (7)
Widowed	268 (17)
Cohabiting	25 (2)
Single	64 (4)
Lives alone	371 (24)
Current employment status	
Full-time paid employment	154 (10)
Part time paid not retired	75 (5)
Part time paid part retired	85 (5)
Sick < 6 months	3 (0)
Sick > 6 months	34 (2)
Fully retired, reached age	761 (49)
Fully retired, early	167 (11)
Fully retired, early, ill health	125 (8)
Carer	8 (1)
Unemployed	9 (1)
Look after home	41 (3)
Perceived financial strain†	
Strain	48 (3)
Be careful	562 (36)
Little difficulty	655 (42)
Comfortable	235 (15)
Area-level deprivation (1=most deprived)‡: median (IQR)	12279.00 (6143.75 to 19489)

Figures are numbers and % of participants unless otherwise stated; IQR Interquartile range
† From {{17446 Thomas, R. 1999}}

Appendix 7 cont: General health characteristics of 1563 eligible participants

Characteristic	
Self-rated health	
Excellent	50 (3.2)
Very good	340 (21.8)
Good	690 (44)
Fair	384 (25)
Poor	86 (6)
Self-report comorbidity:	
Previous fracture: hip	45 (3)
Previous fracture: wrist	227 (15)
Previous fracture: other	510 (33)
Chest problems	350 (22)
Heart problems	348 (22)
Deafness	324 (21)
Eyesight problems	376 (24)
High blood pressure	728 (47)
Diabetes	228 (15)
Previous stroke	77 (5)
Cancer	81 (5)
Liver disease	21 (1)
Kidney disease	37 (2)
Problems with circulation to legs	463 (30)
Falls	239 (15)
Memory	538 (34)
Cough	329 (21)
Breathlessness	620 (40)
Dizziness	409 (26)
Weakness in an arm/leg	411 (26)
Shaking in hands	128 (8)
Swelling ankles/feet	485 (31)
HAD (0-21): median (IQR)	
Anxiety/Depression	5 (3-8)/ 3 (1-6)

Appendix 7 cont: Joint pain characteristics of 1563 eligible participants

Characteristic	Number (%)
Hand pain in past 12 months	
Persistent hand pain (90+)	981 (63)
Average pain intensity (0-10)	246 (16)
Mild (0-3)	988 (63)
Moderate (4-6)	294 (19)
Severe (7-10)	217 (14)
Hip pain in past 12 months	
Persistent hip pain	852 (55)
Average pain intensity (0-10)	192 (12)
Mild (0-3)	967 (62)
Moderate (4-6)	278 (18)
Severe (7-10)	237 (15)
Knee pain in past 12 months	
Persistent knee pain	1113 (71)
Average pain intensity (0-10)	339 (22)
Mild (0-3)	788 (50)
Moderate (4-6)	357 (23)
Severe (7-10)	359 (23)
Pattern of joint pain involvement:	
Hand only	206 (13)
Hip only	129 (8)
Knee only	242 (15)
Hand + Hip	107 (7)
Hand + Knee	249 (16)
Hip + Knee	193 (12)
Hand + Hip + Knee	414 (26)

Appendix 7 cont: pain coping, self-management and healthcare utilisation among 1563 eligible participants

Characteristic	Number (%)
1-item Coping Strategies Questionnaire (0-6): median, IQR	
CSQ 1: Distraction	0 (0-3)
CSQ 2: Reinterpreting pain sensation	0 (0-3)
CSQ 3: Catastrophising	0 (0-2)
CSQ 4: Ignoring pain	2 (0-4)
CSQ 5: Praying and hoping	0 (0-2)
CSQ 6: Coping self-statements	4 (2-6)
CSQ 7: Increased behavioural activities	3 (0-5)

Appendix 8a: Choice questionnaire: the decision to consult the general practitioner for joint pain

KEELE
UNIVERSITY

Arthritis Research UK
Providing answers today and tomorrow

NHS
National Institute for Health Research

Arthritis Research UK Primary Care Centre

CHOICE QUESTIONNAIRE: THE DECISION TO CONSULT THE GENERAL PRACTITIONER FOR JOINT PAIN

REC Number: 10/H1203/63
Version 4, Date: 13/10/2010

Introduction

Thank you for helping with this study. The aim of this questionnaire is to understand what leads people with joint pain to consult a general practitioner (GP) about it. **Even if you don't see a GP very often, we would be very grateful if you would take part as it is important that we hear from a wide variety of people.**

We would like you to consider the importance of some factors that may influence your decision to consult the GP for your joint pain. The factors are things that people with joint problems have told us they might consider when deciding whether to seek medical attention for a joint problem. These include the symptoms you are experiencing at the time and the amount of disruption it is causing to your everyday life.

It is important to remember that t**hese situations are imaginary**. We ask that you try to 'put yourself in each situation', and choose the one that would more likely lead you to consult the general practitioner.

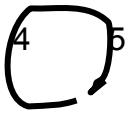
There are no right or wrong answers. We are interested in your own personal preferences based on how you might feel in each situation.

Instructions for this questionnaire

Please answer **all** of the questions.

The questions can be answered by putting a cross in a box like this: ☒

or circling a number like this: 3 4 5 6



Please write in BLOCK CAPITALS where appropriate

When you have finished, please check that you have answered all of the questions and then return the questionnaire in the envelope enclosed. You do not need a stamp.

The answers you give in the questionnaire will be treated in the strictest confidence.

Whether you take part in this research or not, your right to use health services at your practice or elsewhere will not be affected.

If you need help when completing the questionnaire or have any questions please contact our Study Co-ordinator, Domenica Coxon, on **01782 734707.**

Thank you again for your help with this research study.

Section A
Factors that may influence your decision to consult the GP for joint pain

We asked people with joint problems which things might influence their decision to go and see the GP about it. Below is a list of some of the things they came up with. Please read this list and then move on to question 1 on the next page.

a. The type of joint pain you are experiencing at the time

Is it the 'usual aches and pains' or am I getting a different kind of pain, for example, short episodes of unpredictable pain?

b. The level of disruption that the joint pain is causing to your everyday life

How much is the joint problem interfering with your usual activities? Joint problems can disrupt many aspects of everyday living including sleep, daily activities (getting about, looking after the house), family responsibilities, social and leisure activities, and relationships.

c. Other physical health problems that may be bothering you at the time

Have I got any other physical health problems at the moment? These could range from short-lasting minor complaints (e.g. a bad cold, headache) to long-term major health conditions (e.g. diabetes, heart problems).

d. The sorts of investigations the GP is likely to offer you

Will the GP consider giving me a physical examination? Will he/she order appropriate X rays and blood tests?

e. The sorts of treatments the GP is likely to offer you

Will the GP arrange for me to be followed up by a practice nurse or referral to a physiotherapist? Perhaps there is a new promising treatment available?

f. The GP's attitude to your joint problem

Am I going to be told that it's just 'part of the normal ageing process'?

1. Please consider the factors below and circle <u>ONE</u> number from 0 (least important) to 10 (most important) to indicate how important they are in your decision to consult the GP for your joint pain.
(Please circle a number for EVERY factor from a - f)

a. The **type of joint pain** you are experiencing at the time?

0	1	2	3	4	5	6	7	8	9	10
not at all important										extremely important

b. The **level of disruption** the joint problem is causing to your everyday life?

0	1	2	3	4	5	6	7	8	9	10
not at all important										extremely important

c. **Other physical health problems** that may be bothering you at the time?

0	1	2	3	4	5	6	7	8	9	10
not at all important										extremely important

d. The sorts of **investigations** the GP is likely to offer you?

0	1	2	3	4	5	6	7	8	9	10
not at all important										extremely important

e. The sorts of **treatments** the GP is likely to offer you?

0	1	2	3	4	5	6	7	8	9	10
not at all important										extremely important

f. The **GP's attitude** to your joint problem?

0	1	2	3	4	5	6	7	8	9	10
not at all important										extremely important

Section B
Choice Tasks

For each of the following choice tasks numbered 1-10 in this following section, please carefully read the two imaginary situations (A and B) and choose which ONE (A or B) would MORE LIKELY LEAD YOU TO CONSULT A GP IN YOUR PRACTICE FOR YOUR JOINT PAIN by ticking the relevant box beneath.

Please choose one of the OVERALL situations (A or B) even though neither may seem ideal.

At times some of these situations may overlap with each other and may seem to be repetitive for you, but none of these choices are exactly the same and **it is important to us that you consider each choice set separately.**

It is entirely up to you how you make your choices. People have different ways of filling out the choice questionnaire. Some people will look at each situation as a whole and try to compare the whole situation (A) against situation (B), others will compare the factors between situations.

Choice Task 1

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- IF the pain is disrupting **most** of your everyday life
- IF you are experiencing **no other** physical health problems
- IF the GP prescribes pain relief and gives **verbal** advice about your condition

Imaginary Situation B

- IF the pain is disrupting **some** of your everyday life
- IF you are experiencing **other minor** physical health problems
- IF the GP prescribes pain relief, gives **written** advice about your condition and arranges follow-up with a practice nurse and physiotherapy referral

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

(Please remember that these situations are imaginary. We ask that you try to 'put yourself in each situation', and choose the one that would more likely lead you to consult)

Choice Task 2

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

Imaginary Situation A

- IF the GP **investigates with appropriate X rays and blood tests** as well as ask about your symptoms and their effect on your day-to-day life and conduct a thorough physical examination of the joints
- IF the pain is disrupting **most** of your everyday life
- IF you are experiencing **short episodes of more severe, often unpredictable pain**

Imaginary Situation B

- IF the GP **conducts a thorough physical examination of the joints** as well as ask about your symptoms and their effect on your day-to-day life
- IF the pain is **not** disrupting your everyday life
- IF you are experiencing a **dull aching pain, which is there most of the time**

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

Choice Task 3

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- The GP regards your joint pain **as a legitimate health problem that requires treatment**
- You are experiencing **other minor** physical health problems
- You are experiencing a **dull aching pain, which is there most of the time**

Imaginary Situation B

- The GP regards your joint pain **as part of the normal ageing process that one just has to accept**
- You are experiencing **other major** physical health problems
- You are experiencing **short episodes of more severe, often unpredictable pain**

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

Choice Task 4

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- The GP regards your joint pain as **part of the normal ageing process that one just has to accept**
- The GP **investigates with appropriate x rays and blood tests** as well as ask about your symptoms and their effect on your day-to-day life and conduct a thorough physical examination of the joints
- The GP prescribes pain relief and gives **verbal** advice about your condition

Imaginary Situation B

- The GP regards your joint pain as a **legitimate health problem that requires treatment**
- The GP **asks about your symptoms and their effect on your day-to-day life**
- The GP **offers a promising new treatment** as well as prescribes pain relief, gives written advice about your condition and arranges follow-up with a practice nurse and physiotherapy referral

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

Choice Task 5

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- The GP regards your joint pain as a **legitimate health problem that requires treatment**
- The pain is disrupting **some** of your everyday life
- You are experiencing **short episodes of more severe, often unpredictable pain**

Imaginary Situation B

- The GP regards your joint pain as **part of the normal ageing process that one just has to accept**
- The pain is **not** disrupting your everyday life
- You are experiencing a **dull aching pain, which is there most of the time**

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A ☐

Situation B ☐

Choice Task 6

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- You are experiencing **no other** physical health problems
- The GP **offers a promising new treatment** as well as prescribe pain relief, give written advice about your condition and arrange follow-up with a practice nurse and physiotherapy referral
- You are experiencing **short episodes of more severe, often unpredictable pain**

Imaginary Situation B

- You are experiencing **other major** physical health problems
- The GP prescribes pain relief, gives **written** advice about your condition and arranges follow-up with a practice nurse and physiotherapy referral
- You are experiencing a **dull aching pain, which is there most of the time**

I would be more likely to consult the GP about my joint pain IF I was in…

Situation A	Situation B
☐	☐

Choice Task 7

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- The GP regards your joint pain **as a legitimate health problem that requires treatment**
- The GP **conducts a thorough physical examination of the joints** as well as ask about your symptoms and their effect on your day-to-day life
- You are experiencing **no other** physical health problems

Imaginary Situation B

- The GP regards your joint pain as **part of the normal ageing process that one just has to accept**
- The GP **asks about your symptoms and their effect on your day-to-day life**
- You are experiencing **other minor** physical health problems

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

Choice Task 8

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- The GP **conducts a thorough physical examination of the joints** as well as ask about your symptoms and their effect on your day-to-day life
- The pain is disrupting **some** of your everyday life
- The GP **offers a promising new treatment** as well as prescribe pain relief, give written advice about your condition and arrange follow-up with a practice nurse and physiotherapy referral

Imaginary Situation B

- The GP **asks about your symptoms and their effect on your day-to-day life**
- The pain is disrupting **most** of your everyday life
- The GP prescribes pain relief, gives **written** advice about your condition and arranges follow-up with a practice nurse and physiotherapy

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

Choice Task 9

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- The GP **investigates with appropriate x rays and blood tests** as well as ask about your symptoms and their effect on your day-to-day life and conduct a thorough physical examination of the joints
- You are experiencing **other major** physical health problems
- You are experiencing a **dull aching pain, which is there most of the time**

Imaginary Situation B

- The GP **conducts a thorough physical examination of the joints** as well as ask about your symptoms and their effect on your day-to-day life
- You are experiencing **other minor** physical health problems
- You are experiencing **short episodes of more severe, often unpredictable pain**

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

(Please remember that these situations are imaginary. We ask that you try to 'put yourself in each situation', and choose the one that would more likely lead you to consult)

Choice Task 10

Imagine that your joint problem and your general health are changeable. Put yourself in each of the situations below, then choose the one that you think would more likely lead you to consult the general practitioner about your joint pain. Please read each situation carefully and, when you have made your choice, **PUT A CROSS IN ONE BOX ONLY**

IF……..

Imaginary Situation A

- The GP regards your joint pain as **part of the normal ageing process that one just has to accept**
- The pain is disrupting **most** of your everyday life
- The GP **offers a promising new treatment** as well as prescribe pain relief, give written advice about your condition and arrange follow-up with a practice nurse and physiotherapy referral

Imaginary Situation B

- The GP regards your joint pain **as a legitimate health problem that requires treatment**
- The pain is **not** disrupting your everyday life
- The GP prescribes pain relief and gives **verbal** advice about your condition

I would be more likely to consult the GP about the joint problem IF I was in…

Situation A	Situation B
☐	☐

Now, please answer the following question:

How difficult was it for you to imagine these situations overall?
(Please put a cross in one box only)

Not at all hard…… ☐

A little hard …… ☐

Quite hard …… ☐

Very hard …… ☐

Extremely hard…… ☐

12. **We would like to invite any other comments you may have on…**

(a) **the choice tasks, for example, the difficulty of the tasks, layout etc?**

(b) **any other comments on deciding whether to see the GP about your joint pain, for example, are there any other vital factors that we have missed out of this questionnaire?**

Section C
Joint pain

The following questions are about **your joint pain**.

1. Have you had any pain in the last year in or around the **HIP**?
(Please place a cross in one box only)

Yes….. ☐ No….. ☐

2. Have you had any pain in the last year in or around the **KNEE**?
(Please place a cross in one box only)

Yes….. ☐ No….. ☐

3. Have you had any pain in your **HANDS**, including in your fingers and thumbs, over the **last year**?
(Please place a cross in one box only)

Yes….. ☐ No….. ☐

4. How long is it since you had a whole month without any joint pain?
(Please put a cross in one box only)

Less than 3 months	3-6 months	7-12 months	1-2 years	3-5 years	6-10 years	More than 10 years
☐	☐	☐	☐	☐	☐	☐

5. How long ago did you first start having joint pain?
(Please put a cross in one box only)

Less than one year ago	1-5 years ago	6-10 years ago	More than 10 years ago
☐	☐	☐	☐

6. Have you consulted any of the following for joint pain during the past 12 months?

(Please put a cross in as many boxes as necessary)

a. General practitioner (GP)……..	☐
b. Hospital specialist……	☐
c. Physiotherapist……..	☐
d. Occupational therapist……..	☐
e. Nurse……...	☐
f. Chiropractor……	☐
g. Osteopath……...	☐
h. Acupuncturist……..	☐
i. Homeopath…….	☐
j. Aromatherapist……...	☐
k. Other (please specify) _____	

7. Have you received any of the following for your current joint problem during the past 12 months?

(Please put a cross in as many boxes as necessary)

a. MRI scan……	☐
b. X Ray……...	☐
c. Other (please specify) _____	☐

8. When did you **first go to your GP** about joint pain?
(Please put a cross in one box only)

Never been to GP	Less than one year ago	1-5 years ago	6-10 years ago	More than 10 years ago
☐	☐	☐	☐	☐

9. For the following questions, please circle the number that best corresponds to your views:

a. How much does your joint pain affect your life?

0	1	2	3	4	5	6	7	8	9	10
no effect at all										severely affects my life

b. How long do you think your joint pain will continue?

0	1	2	3	4	5	6	7	8	9	10
a very short time										forever

c. How much control do you feel you have over your joint pain?

0	1	2	3	4	5	6	7	8	9	10
absolutely no control										extreme amount of control

d. How much do you think your GP can help your joint pain?

0	1	2	3	4	5	6	7	8	9	10
not at all										extremely helpful

e. How much do you experience symptoms from your joint pain?

0	1	2	3	4	5	6	7	8	9	10
no symptoms at all										many severe symptoms

f. How much does your joint pain affect you emotionally (e.g. does it make you angry, scared, upset or depressed)?

0	1	2	3	4	5	6	7	8	9	10
not at all emotionally affected										extremely emotionally affected

g. Please list in rank-order the three most important factors that you believe caused <u>your joint pain</u>. *The most important causes for me:-*

(i) _____

(ii) _____

(iii) _____

h. How concerned are you about your joint pain?

0	1	2	3	4	5	6	7	8	9	10
not at all concerned										extremely concerned

i. How well do you feel you understand your joint pain?

0	1	2	3	4	5	6	7	8	9	10
don't understand at all										understand very clearly

Section D
About you

1. What is your date of birth? ☐☐ / ☐☐ / ☐☐

(For example – if you were born on the 5th June 1936, this would be entered as 05/06/36)

2. Are you:

Male……. ☐ Female……. ☐

3. What is your current employment status?
(Please put a cross in one box only)

Employed…….. ☐
Not working due to ill health…… ☐
Retired…… ☐
Unemployed / seeking work…… ☐
Housewife……... ☐
Other…… ☐

Section E
Continuing to help with this study

Thank you for completing this questionnaire

Please ensure that you have read the enclosed information sheet that explains the study in greater detail. There will be further stages to this study, and we hope that people who have taken part so far will be able to help us again. We may want to contact you again as part of this study and we are asking your permission to do this. Giving us permission to contact you again does not mean that you must take part.

Please see page 26 overleaf for the consent form for this questionnaire

Study ID number
Version no. 1

CONSENT FORM

Please read and complete the following consent form, and then sign below.

Please answer each statement by putting a cross in one box on each line

	Yes	**No**
I confirm that I have read and understood the study information sheet and am willing to take part in the study.....	☐	☐
I understand that I can withdraw from the study at any time, and that this will not affect the care I receive in any way.....	☐	☐
I am happy to be contacted again (this does not mean that you must take part in future - you are just agreeing to be contacted again).....	☐	☐

Signed ..... Date .....

Please print your name and address .....

.....

.....

Tel. number .....

Even if you would prefer us not to contact you again about the study, the answers you have given in this questionnaire will still be very important to us.

Please return your questionnaire in the FREEPOST (no stamp needed) envelope provided

If you have any questions, telephone Domenica Coxon on 01782 734707.

Thank you for your help with this research study

Appendix 8b: Questionnaire cover letter

KEELE UNIVERSITY

Arthritis Research UK
Providing answers today and tomorrow

NHS National Institute for Health Research

PRIMARY CARE SCIENCES

ARTHRITIS RESEARCH UK PRIMARY CARE CENTRE

Dear XXXX,

The researchers in the Arthritis Research UK Primary Care Centre at Keele University are writing to you to see if you would be willing to help us with a research study.

Researchers at Keele are trying to find out how people with joint pain in North Staffordshire make the decision to go and visit a general practitioner, to get a better understanding of how this service might be improved. Further details of the study are on the accompanying information sheet.

We are inviting patients aged 50 years and over, who have kindly helped with research before, and who indicated in their last questionnaire that they had experienced joint pain, to take part. We hope you will be able to spare some of your time to complete the enclosed questionnaire. It should take you no longer than about 60 minutes. We are very interested in your reply, even if you have not recently consulted your general practitioner for your joint pain.

All of your answers will be dealt with in strict confidence. We can also assure you that whether you answer the questionnaire or not, this will not in any way affect the care you receive from this practice or elsewhere.

We would be very grateful if you would return the questionnaire in the next two weeks in the envelope provided (no stamp is needed). A short while after this date, we will send a reminder to people whose questionnaire we have not received. If you would like to know more about this study, please contact **Domenica Coxon** at Keele University on 01782 734707.

Thank you very much for your help with this research study.

Yours sincerely

Dr George Peat
Senior Lecturer in Clinical Epidemiology,
Arthritis Research UK Primary Care Centre,
Keele University
Keele, Staffordshire
ST5 5BG
UK

Appendix 8c: Patient information sheet

KEELE UNIVERSITY

Arthritis Research UK
Providing answers today and tomorrow

NHS National Institute for Health Research

PRIMARY CARE SCIENCES

ARTHRITIS RESEARCH UK PRIMARY CARE CENTRE

CHOICE QUESTIONNAIRE: THE DECISION TO CONSULT THE GENERAL PRACTITIONER FOR JOINT PAIN

You are being invited to take part in a research study. Before you decide whether to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully.

What is the purpose of the study?
Joint pain is a very common problem. The general practitioner (GP) is often the first port of call. However, we know that most people do not go to the GP about it, even though some may have quite severe problems. We want to understand what influences people's decisions to consult the GP about joint pain. It may suggest how services could be improved.

Why have I been invited?
As a local resident who has previously taken part in research and indicated that they had experienced joint pain, we are interested in some of the things that may influence your decision to consult a GP for joint pain. If you could spare time to fill in the enclosed questionnaire you would provide information that will be of great benefit for this study. We would like you to take part and fill in the questionnaire even if you do not currently have joint pain and even if you have not been to the GP recently. The information you give us will be stored and will help the development of further studies.

Do I have to take part?
Whether or not you take part in this research is voluntary. If you do decide to take part, you are free to withdraw at any time without giving a reason. A decision to withdraw, or a decision not to take part, will not affect your right to access health services at your practice or elsewhere.

How long will it take?
We think it will take you about 60 minutes to fill in the questionnaire.

Future contact
In the future, we may contact you again to ask you further questions about this research. We ask for your permission to contact you again on the last page of the questionnaire. If you agree to be contacted again, this does not mean that you must take part in future; you are only agreeing to be contacted again.

What are the possible benefits of taking part?
Although any direct benefit to you is unlikely, what we learn from the study will help people with joint pain in the future. This research will benefit people with joint pain as we learn about the types of health factors and service factors that are influencing the

decision to consult. Only by gaining your views can we learn about the different factors that are involved. We will be able to make better decisions on the most important types of services and how many people will require services in the future.

Will my taking part in this study be kept confidential?
The answers you give in the questionnaire will be dealt with in strictest confidence. Each person who responds to the questionnaire will be given a code number, so the data from the study will not have any identifiable names and addresses, and cannot be traced back to you. On this basis, the data may be used in other research studies. The questionnaires will be stored without identifiable names and addresses for twenty years in accordance with the Medical Research Council guidelines. Beyond this date records will be maintained if the study is still ongoing. The questionnaires will be stored in a secure place.

What will happen if I don't want to carry on with this study?
You can withdraw from this study by telephoning us on 01782 734707. Withdrawing means that we would no longer contact you directly, but we would still keep and use the information you have provided up to the point of your withdrawal. If you contact us to withdraw from the study, and you have consented to medical record review, we will check whether you also want us to stop reviewing your medical records.

What will happen to the results of the research study?
A summary of the findings will be sent to you and a poster will be available for you to look at in the surgery. If you would like any other information after seeing these we will be happy to help.

Who is funding and organising the research?
The Arthritis Research UK is funding the research, which is organised by the Arthritis Research UK Primary Care Centre at Keele University.

Who has reviewed the study?
The Staffordshire Research Ethics Committee has reviewed this study (Research Ethics Committee Reference Number: 10/H1203/63)

Contact for further information
If you have any questions, or would like further information, about this study please contact our Study Co-ordinator, Domenica Coxon on 01782 734707.

If you have any questions or concerns about taking part in this research you can also contact the Patient Advice and Liaison Service (PALS). Your local PALS office free phone number for Stoke on Trent is 0800 783 2865 and for North Staffordshire is 0800 389 8832.

Thank you for taking time to read this information leaflet.

Appendix 8d: Reminder postcard

CHOICE QUESTIONNAIRE

We are writing to remind you of a questionnaire we recently sent to you asking about your health. This is part of a study at Keele University. We are still interested in hearing from you. We would be grateful if you could fill in the questionnaire and return it in the reply paid envelope as soon as you can. **Your answers are strictly confidential**. If you have returned the questionnaire in the last few days, please ignore this postcard and we apologise for troubling you. If you have any questions about this study please telephone the study co-ordinator Domenica Coxon on **01782 734707.**

Thank you very much for your help with this research.

Appendix 8e: Repeat cover letter

KEELE UNIVERSITY

Arthritis Research UK
Providing answers today and tomorrow

NHS National Institute for Health Research

PRIMARY CARE SCIENCES

ARTHRITIS RESEARCH UK PRIMARY CARE CENTRE

Dear XXXX,

We are writing to remind you of the study that researchers at the Arthritis Research UK Primary Care Centre at Keele University are undertaking, where we are interested in finding out about how adults make the decision to consult a general practitioner about their joint pain.

So far we don't seem to have received a reply from you. We are therefore sending you a second questionnaire in case you mislaid the first one. We would be very grateful if you could spare a few minutes of your time to complete this questionnaire which asks you how you make the decision to consult a general practitioner about your joint pain. The questionnaire should take you about 60 minutes to complete.

All your answers will be treated in the strictest confidence and used only for the purpose of this research study. If you have any difficulty filling in the questionnaire please contact **Domenica Coxon** on 01782 734707.

We would be grateful if you could return the completed questionnaire in the next two weeks in the envelope provided. **You do not need a stamp**.

We hope that you will take part in this study. All of your answers will be dealt with in strict confidence. We can also assure you that whether or not you answer the questionnaire will not in any way affect the care you receive from this practice or elsewhere. If you have returned the questionnaire in the last few days, please ignore this letter and we apologise for troubling you again.

Thank you very much for your help with this research study.
Yours sincerely

Dr George Peat
Senior Lecturer in Clinical Epidemiology,
Arthritis Research UK Primary Care Centre,
Keele University
Keele, Staffordshire
ST5 5BG, UK

Appendix 9: Multinomial logit explained

A partworth utility is a measure of relative desirability or value (Orme 2006). From the multinomial logit models every attribute level is assigned a partworth utility. In the context of this study the higher the utility the greater the relative impact that attribute level has on the decision to consult the general practitioner for joint pain. (Orme 2010).

When using MNL, the raw utilities are zero-centred within each attribute, for example:

Attributes and levels				Partworth utility
2	Level of disruption to everyday life	2.1	None	-0.65
		2.2	Some	0.20
		2.3	Most	0.45

The above example shows that respondents are more likely to consult the general practitioner for joint pain when they have most pain disruption to their everyday life than when they have lower levels of pain disruption (Orme, 2010).

The aggregate logit model combines all respondents' responses and estimates a single set of effects (partworth utilities) based on the total sample. It is done such that the choices actually made are predicted according to the logit rule.

For example, consider 2 alternatives (A and B) in a choice set:

$PA = \exp(UA) / [\exp(UA)+\exp(UB]$

Where *PA*= probability of choosing alternative A and exp means exponentiate or antilog and UA= total utility for alternative A.

So, for example, if product A had a total utility of 2.0; and Product B had total utility of 3.0, the likelihood of choosing rather than B is

PA= exp(2.0) / [exp(2.0)+exp(3.0)]

= 7.39 / (7.39+20.09)

= 26.9%

The probability of choosing scenario B is PB= 1.0 –0.269 = 73.1% (Orme 2006).

MNL uses an iterative procedure to find the maximum likelihood solution for fitting a MNL model to the data. Maximum likelihood estimation estimates partworth utilities with the aim to maximise the fit to respondents' observed choices. Likelihood is a measure of fit that indicates the probability of the observed choices given the estimated partworth utilities (Orme 2006)

Appendix 10: Highest 50 combinations of scenarios (2*3*3*3*3*2=324) ranked in descending order of overall utility

Combination	Att1	Att2	Att3	Att4	Att5	Att6	Uatt1	Uatt2	Uatt3	Uatt4	Uatt5	Uatt6	Sum	Rank
324	2	3	3	3	3	2	0.08121	0.44714	0.23335	0.21143	0.19388	0.43179	1.5988	1
318	2	3	3	2	3	2	0.08121	0.44714	0.23335	0.21143	0.06594	0.43179	1.47086	2
322	2	3	3	3	2	2	0.08121	0.44714	0.23335	0.0618	0.19388	0.43179	1.44917	3
162	1	3	3	3	3	2	-0.08121	0.44714	0.23335	0.21143	0.19388	0.43179	1.43638	4
306	2	3	2	3	3	2	0.08121	0.44714	-0.0107	0.21143	0.19388	0.43179	1.35475	5
270	2	2	3	3	3	2	0.08121	0.20114	0.23335	0.21143	0.19388	0.43179	1.3528	6
316	2	3	3	2	2	2	0.08121	0.44714	0.23335	0.0618	0.06594	0.43179	1.32123	7
156	1	3	3	2	3	2	-0.08121	0.44714	0.23335	0.21143	0.06594	0.43179	1.30844	8
160	1	3	3	3	2	2	-0.08121	0.44714	0.23335	0.0618	0.19388	0.43179	1.28675	9
300	2	3	2	2	3	2	0.08121	0.44714	-0.0107	0.21143	0.06594	0.43179	1.22681	10
264	2	2	3	2	3	2	0.08121	0.20114	0.23335	0.21143	0.06594	0.43179	1.22486	11
304	2	3	2	3	2	2	0.08121	0.44714	-0.0107	0.0618	0.19388	0.43179	1.20512	12
268	2	2	3	3	2	2	0.08121	0.20114	0.23335	0.0618	0.19388	0.43179	1.20317	13
144	1	3	2	3	3	2	-0.08121	0.44714	-0.0107	0.21143	0.19388	0.43179	1.19233	14
108	1	2	3	3	3	2	-0.08121	0.20114	0.23335	0.21143	0.19388	0.43179	1.19038	15
154	1	3	3	2	2	2	-0.08121	0.44714	0.23335	0.0618	0.06594	0.43179	1.15881	16
312	2	3	3	1	3	2	0.08121	0.44714	0.23335	0.21143	-0.25982	0.43179	1.1451	17
288	2	3	1	3	3	2	0.08121	0.44714	-0.22264	0.21143	0.19388	0.43179	1.14281	18
320	2	3	3	3	1	2	0.08121	0.44714	0.23335	-0.27323	0.19388	0.43179	1.11414	19
252	2	2	2	3	3	2	0.08121	0.20114	-0.0107	0.21143	0.19388	0.43179	1.10875	20
298	2	3	2	2	2	2	0.08121	0.44714	-0.0107	0.0618	0.06594	0.43179	1.07718	21
262	2	2	3	2	2	2	0.08121	0.20114	0.23335	0.0618	0.06594	0.43179	1.07523	22
138	1	3	2	2	3	2	-0.08121	0.44714	-0.0107	0.21143	0.06594	0.43179	1.06439	23
102	1	2	3	2	3	2	-0.08121	0.20114	0.23335	0.21143	0.06594	0.43179	1.06244	24
142	1	3	2	3	2	2	-0.08121	0.44714	-0.0107	0.0618	0.19388	0.43179	1.0427	25

Appendix 10 cont: Highest 50 combinations of scenarios (2*3*3*3*3*2=324) ranked in descending order of overall utility

Combination	Att1	Att2	Att3	Att4	Att5	Att6	Uatt1	Uatt2	Uatt3	Uatt4	Uatt5	Uatt6	Sum	Rank
106	1	2	3	3	2	2	-0.08121	0.20114	0.23335	0.0618	0.19388	0.43179	1.04075	26
282	2	3	1	2	3	2	0.08121	0.44714	-0.22264	0.21143	0.06594	0.43179	1.01487	27
310	2	3	3	1	2	2	0.08121	0.44714	0.23335	0.0618	-0.25982	0.43179	0.99547	28
286	2	3	1	3	2	2	0.08121	0.44714	-0.22264	0.0618	0.19388	0.43179	0.99318	29
314	2	3	3	2	1	2	0.08121	0.44714	0.23335	-0.27323	0.06594	0.43179	0.9862	30
150	1	3	3	1	3	2	-0.08121	0.44714	0.23335	0.21143	-0.25982	0.43179	0.98268	31
246	2	2	2	2	3	2	0.08121	0.20114	-0.0107	0.21143	0.06594	0.43179	0.98081	32
126	1	3	1	3	3	2	-0.08121	0.44714	-0.22264	0.21143	0.19388	0.43179	0.98039	33
250	2	2	2	3	2	2	0.08121	0.20114	-0.0107	0.0618	0.19388	0.43179	0.95912	34
158	1	3	3	3	1	2	-0.08121	0.44714	0.23335	-0.27323	0.19388	0.43179	0.95172	35
90	1	2	2	3	3	2	-0.08121	0.20114	-0.0107	0.21143	0.19388	0.43179	0.94633	36
136	1	3	2	2	2	2	-0.08121	0.44714	-0.0107	0.0618	0.06594	0.43179	0.91476	37
100	1	2	3	2	2	2	-0.08121	0.20114	0.23335	0.0618	0.06594	0.43179	0.91281	38
294	2	3	2	1	3	2	0.08121	0.44714	-0.0107	0.21143	-0.25982	0.43179	0.90105	39
258	2	2	3	1	3	2	0.08121	0.20114	0.23335	0.21143	-0.25982	0.43179	0.8991	40
234	2	2	1	3	3	2	0.08121	0.20114	-0.22264	0.21143	0.19388	0.43179	0.89681	41
302	2	3	2	3	1	2	0.08121	0.44714	-0.0107	-0.27323	0.19388	0.43179	0.87009	42
266	2	2	3	3	1	2	0.08121	0.20114	0.23335	-0.27323	0.19388	0.43179	0.86814	43
280	2	3	1	2	2	2	0.08121	0.44714	-0.22264	0.0618	0.06594	0.43179	0.86524	44
120	1	3	1	2	3	2	-0.08121	0.44714	-0.22264	0.21143	0.06594	0.43179	0.85245	45
148	1	3	3	1	2	2	-0.08121	0.44714	0.23335	0.0618	-0.25982	0.43179	0.83305	46
244	2	2	2	2	2	2	0.08121	0.20114	-0.0107	0.0618	0.06594	0.43179	0.83118	47
124	1	3	1	3	2	2	-0.08121	0.44714	-0.22264	0.0618	0.19388	0.43179	0.83076	48
152	1	3	3	2	1	2	-0.08121	0.44714	0.23335	-0.27323	0.06594	0.43179	0.82378	49
84	1	2	2	2	3	2	-0.08121	0.20114	-0.0107	0.21143	0.06594	0.43179	0.81839	50

Appendix 11a: Joint pain and healthcare utilisation characteristics of 250 randomly selected conjoint respondents, stratified by subgroups 1-3

	Subgroup 1: Strong	Subgroup 2: Weak	Subgroup 3: Intermediate	*p* value
N	114	63	73	
Hip pain in past 12 months	75 (66)	30 (48)	48 (66)	.045
Knee pain in past 12 months	94 (82)	45 (71)	50 (68)	.064
Hand pain in past 12 months	83 (73)	43 (68)	46 (63)	.368
Average Pain intensity (0-10) Median (IQR)				
Hip pain intensity	2 (0,5)	1 (0,5)	2 (0,6)	.582
Knee pain intensity	4 (0,7)	3 (0,7)	4 (1,7)	.914
Hand pain intensity	1.5 (0,4)	3 (0,6)	2 (0,4)	.031
How long since pain first started:				
<1 year ago	3 (3)	0 (<0)	3 (4)	.241
1-5 years ago	29 (25)	24 (38)	21 (29)	
6+ years ago	82 (72)	38 (60)	49 (67)	
How long since month without joint pain†:				
<3 months	36 (32)	32 (51)	27 (37)	.080
3-12 months	21 (18)	9 (14)	6 (8)	
1-5 years	23 (20)	12 (19)	14 (19)	
6+ years	34 (30)	10 (16)	24 (33)	
First consulted GP for joint pain				
Never been to GP	14 (12)	9 (14)	16 (22)	.042
0-5 years ago	33 (29)	29 (46)	22 (30)	
6+ years ago	66 (58)	24 (38)	35 (48)	
Consulted for joint pain in past 12 months:				
GP	53 (46)	36 (57)	37 (51)	.398
Hospital specialist	22 (19)	8 (13)	15 (21)	.438
Physiotherapist	22 (19)	8 (13)	13 (18)	.531
Occupational Therapist	3 (3)	2 (3)	2 (3)	.978
Nurse	14 (12)	3 (5)	5 (7)	.188
Chiropractor	4 (4)	0 (0)	2 (3)	.336
Osteopath	3 (3)	2 (3)	3 (4)	.855
Acupuncturist	1 (1)	6 (10)	6 (8)	.018
Homeopath	0 (0)	1 (2)	1 (1)	.425
Aromatherapist	3 (3)	0 (1<)	2 (3)	.423

Appendix 11b: Number of self-reported comorbidities, anxiety and depression scores of 250 randomly selected conjoint respondents, stratified by subgroups 1-3

	Subgroup 1: Strong	Subgroup 2: Weak	Subgroup 3: Intermediate	*p* value
N	114	63	73	
Number of self-reported comorbidities (0-22) median (IQR)	4 (2,7)	4.5(2,6)	4 (2,6)	.513
HAD (0-21): Median (IQR) Anxiety Depression	 5 (2,7) 3 (1,5)	 6 (3,8) 3 (1,7)	 5 (2,9) 3 (1,6)	 .323 .921

Appendix 11c: Brief Illness Perceptions Questionnaire and Coping Strategies Questionnaire (CSQ) of 250 randomly selected conjoint respondents, stratified by subgroups 1-3

	Subgroup 1: Strong	Subgroup 2: Weak	Subgroup 3: Intermediate	*p* value
N	114	63	73	
Individual items (0-10):				
BIPQ 1: Consequence	5 (2.5,7.5)	4 (2,5)	5 (2,8)	.157
BIPQ 2: Timeline	10 (8.5, 10)	10 (7,10)	10 (8, 10)	.304
BIPQ 3: Personal control	5 (4,7)	5 (3,8)	5 (4,8)	.751
BIPQ 4: Treatment control	5 (3,7)	5 (2,7)	4 (2,7)	.607
BIPQ 5: Identity	7 (4,8)	5 (3,8)	6 (4,8)	.203
BIPQ 6: Emotional representation	3 (1,5)	3 (1,5)	3 (1,6)	.818
BIPQ 8: Illness concern	5 (3,8)	4 (3,7)	5 (3,8)	.357
BIPQ 9: Coherence	3 (2,5)	5 (2,7)	3 (1,5)	.026
Coping strategies questionnaire (CSQ) Items 1-7 (%, scoring >0 on each item):				
CSQ 1: Distraction	66 (58)	32 (51)	38 (52)	.416
CSQ 2: Reinterpreting pain sensation	62 (54)	34 (54)	43 (59)	.617
CSQ 3: Catastrophising	58 (51)	29 (46)	45 (62)	.122
CSQ 4: Ignoring pain	21 (18)	19 (30)	22 (30)	.099
CSQ 5: Praying and hoping	69 (61)	34 (54)	42 (58)	.666
CSQ 6: Coping self-statements	15 (13)	10 (16)	13 (18)	.660
CSQ 7: Increased behavioural activities	29 (25)	19 (30)	21 (29)	.712